

Access DB#

62595

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

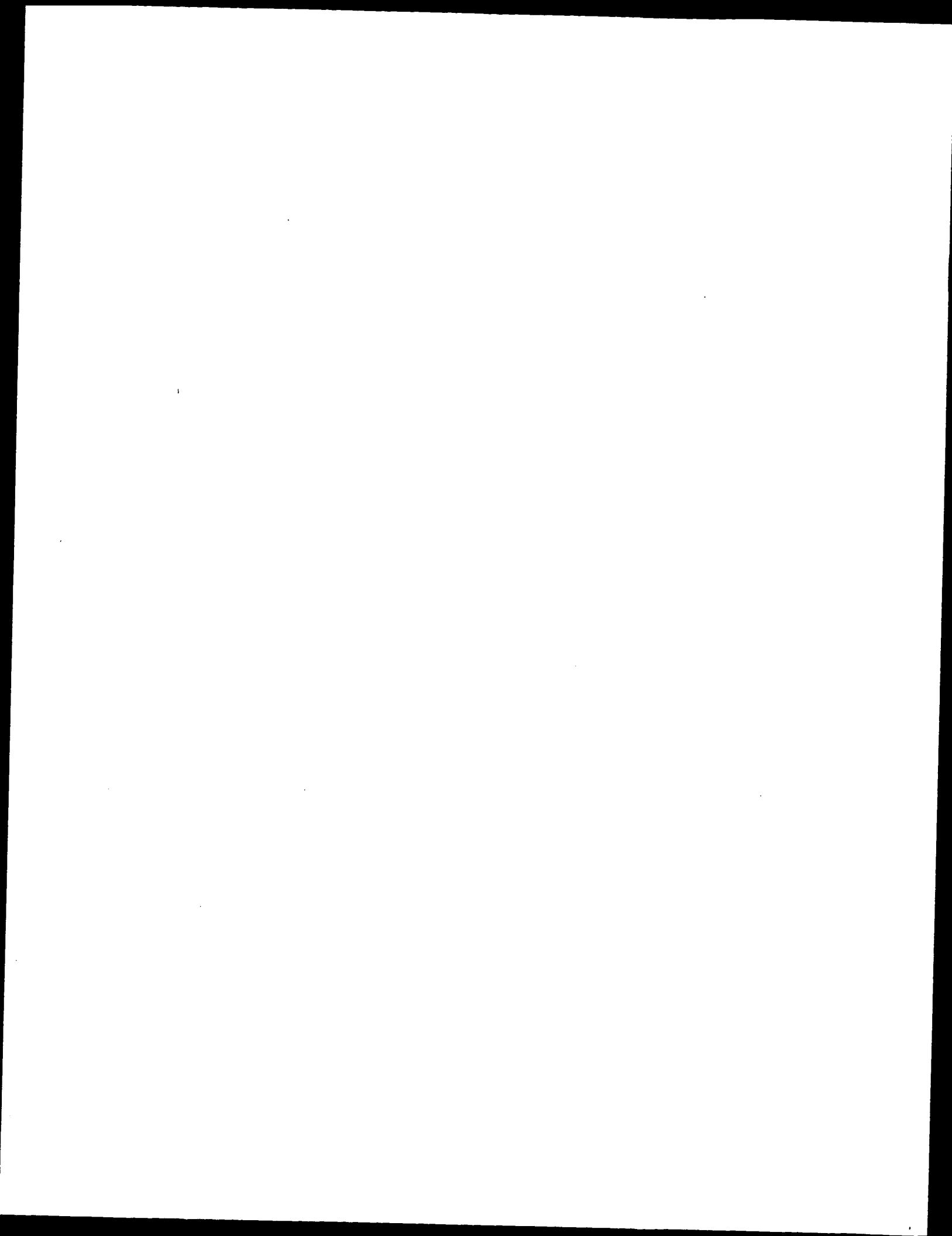
Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval  
Reference Librarian  
Biotechnology & Chemical Library  
CM1 1E07 - 703-308-4498  
jan.delaval@uspto.gov

\*\*\*\*\*  
STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <input checked="" type="checkbox"/>	STN _____
Searcher Phone #: 4498	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: 3/18/02	Bibliographic _____	Dr.Link _____
Date Completed: 3/18/02	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/>
Clerical Prep Time: 10	Patent Family _____	WWW/Internet _____
Online Time: 10	Other _____	Other (specify) _____





A:Gene: clfl  
 C:Keywords: cytokine; glycoprotein  
 F:164/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 118.5; DB 2; Length 203;  
 Best Local Similarity 28.8%; Pred. No. 0.0012;  
 Matches 51; Conservative 21; Mismatches 88; Indels 17; Gaps 5;

40 IOKTYDLTRYLEHQLRSLAGTYLNTYLGPPNEDFNPRL---GAETLPRAVDLEWRS 96  
 27 IROTHSLAHLTKYAEQLLEEVYQGGEPGLGFSPPRLPLAGLSGAPASHAGLPV--- 83  
 97 LNDKLRNTONYEAVSHLCTYLRGLNROA-----TAEILRSIAHFTSLQGLSIGVMA 151  
 84 ---SERLRQDAALSVLPALDLAVRRQAEINPRAPRLRLSLIEDARQVRLGAIVETVL 140  
 152 AALGYPVLP---POLPFGTEPI---WTPGPAHSDFLOKMDFFWLKELQTWLMSAKDFNRL 205  
 141 AALGAAPGPEPVYATLFTANSTAGIFSAKVLGFHVCGLGHWVSRTEDLQGL 197

RESULT 3  
 686182  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: G66182  
 R:Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.; Sun, H.; Tallon, A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: G66182  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-332 <STO>  
 A:Cross-references: GB:AE005172; MID:g7211973; PIDN:AAF0444.1; GSDB:G000141  
 C:Genetics:  
 A:Map position: 1

Query Match 7.5%; Score 92; DB 2; Length 332;  
 Best Local Similarity 26.4%; Pred. No. 0.64;  
 Matches 58; Conservative 24; Mismatches 72; Indels 66; Gaps 13;

23 PAVPALNRTGDPGSPISOKTYDLTRYLEHQLSLAGTYLNTY----- 64  
 113 PSYTAGLMSGP-PRSEF--TYDGPYEQOMESLLOQFIRENPQIRPLRLGLSPVG 169  
 65 LGPFENPDENPRPGAETLPRAVDLEWRSLSNDKRLRLTONEAVSHLCTYLRGLNROA 124  
 170 LGPIRASPOFLOGRVAP-PTSILD-----TSNRKAR-----SDGALAAYRG--RKV 215  
 125 ATTELKRLS-----AHFCTSLQGLSIGVMAALGYPLPQPLP---GTEPTWT 170  
 216 RTTESSSLYSLSGRSWLKNGAHV---GIOPQSGIMR---PLPKPLPVLTLETTSVP 266  
 171 PGFAHSDFLQKMDFFWLKELQTWLMSAKDFNRLKKKKQ 210  
 267 DDPDESSADEKDEDEAVKQL-----SEKDL--LKRHE 298

RESULT 4  
 G02312  
 cardiotrophin-1 - human  
 C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998  
 C:Accession: G02312  
 R:Wood, W.T.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: H01035  
 A:Accession: G02312  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-201 <MOO>  
 A:Cross-references: EMBL:U43030; MID:g1151149; PID:g1151150  
 C:Genetics:  
 A:Gene: GDB:CTF1; CT-1  
 A:Cross-references: GDB:567078  
 A:Map position: 1p22-1p22

Query Match 7.5%; Score 91.5; DB 2; Length 201;  
 Best Local Similarity 25.9%; Pred. No. 0.38;  
 Matches 45; Conservative 24; Mismatches 92; Indels 13; Gaps 5;

40 IOKTYDLTRYLEHQLRSLAGTYLNTYLGPPNEDFNPRLGAEILPRAVDLEWRSLSND 99  
 27 IROTHSLAHLTKYAEQLLEEVYQGGDPRLGFSPPRLPVAGL-SATAPSIAGLPVHE 85  
 100 KIRLTQNYEAVSHLCTYLRGLNROA-----TAEILRSIAHFTSLQGLSIGVMAAL 154  
 86 RLRL--DAALALAPLLDLAVCRQAEINPRAPRLRLSLIEDARQVRLGAIVETVL 143  
 155 GYPLPQPLPCTE---TWTPGPAHSDFLOKMDFFWLKELQTWLMSAKDFNRL 205  
 144 G--AANRGAEPRAVPATASASATGVFPAKVLGLRVCGLREWLSTEDLQGL 195

RESULT 5  
 148378  
 hairless protein - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 01-Nov-1999  
 C:Accession: I48378  
 R:Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; St, S.; Stoye, J.P.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994  
 A:Title: Structure and expression of the hairless gene of m  
 A:Reference number: I48378; MUID:94329587  
 A:Accession: I48378  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1182 <RES>  
 A:Cross-references: EMBL:Z32675; MID:g531706; PIDN:CAA83987.1; PID:g531707

Query Match 7.5%; Score 91.5; DB 2; Length 1182;  
 Best Local Similarity 23.7%; Pred. No. 3.5;  
 Matches 70; Conservative 25; Mismatches 99; Indels 101; Gaps 17;

3 LRAGSGMGLACLVTLVHLPAVPAINRTGDPG-----PGP----- 38  
 813 LRAGS--GLRKL-----SLPLSPVTRLSPGALLWLQEPKPHGFLRQELMKQGPV 865  
 39 ---SIOKTYDLTRYLEHQLRSLAGTYLNTYLGPP-----FNEPPEPNPRLG 80  
 866 LVSGIQKTLRLSIMOMELGTLCGOVOTLTALGPQPTNDSTAFWBCFSHPETRP---- 921  
 81 AETLPRAVDLEWRSLSNDK-LRLTON-----YEAVSHLCTYLRGLNROAAT 126  
 922 --KLDEGSV-LLHRLTGDKDASRVONLASSLPPEYCAHQKINLASYLPLGLTTLPLE 978  
 127 AEL-----RSLAHFTSLQGLSIGVMAALGYPLPQPLPCTEPTWTGPAH 175  
 979 POLMAAYGVNSHRLGHTKNCVEVSDLSITLVAEQL-----PMWY--RAQ 1024  
 176 SDFLOKMD--DFWLKELQTWIMR--SAKDFNRLKKKKMP--PAAAVTILALGCHG 224  
 1025 KDFLSGLDGGSLWSPGQSTVWVHVFRAQDAQXIRRELQWVCPAGACTLPEPAG 1079



```
Best Local Similarity    22.8%;   Pred. No. 6.2;          105; Gaps      147
Matches        61; Conservative    32; Mismatches              69;             105; Gaps      147
```

749. score 91; DB 2; Length 195;

Score 89.5; DB 2; Length 1313;

1014 TOWDWR--NRCLNSFKTSASAT 1036

A;Gene: PA0434  
hypothetical protein HI1680

070 10PSP-----RS1,KDAFERLRQOLT-----TSLLEFRHG 400

## RESULT

C:Accession: T00247  
 R:Matsumoto, K.; Ishii, N.; Yoshida, S.; Shiosaka, S.; Watanaka, A.; Tohyama, M.  
 submitted to the EMBL Data Library, March 1998  
 A:Description: Molecular cloning and distinct developmental expression pattern of splice  
 A:Reference number: 214130  
 A:Accession: T00247  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-955 <MAT>  
 A:Cross-references: EMBL:AB012266; NID:d1227741; PIDN:BAA32791.1; PID:d1033757  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: wiz

Query Match  
 Best Local Similarity 7.1%; Score 87; DB 2; Length 955;  
 Matches 35; Conservative 19; Mismatches 48; Indels 24; Gaps 7;

QY 22 LPVPAALNRGTDGPGPSIQKTYDTRYLEHQLRLAG---TYLNYLGP-----PENEPPD 73  
 Db 615 LPUSPLASRPKPGAGPT-----QVPR--ELSLSPITSKPSAASYLGVPATKRPLQEDR 667  
 QY 74 FNPRLGAETLPRAVDLEWMSLNDKLRITONTYEAASHLLCYLRGL---NRQATATELR 130  
 Db 668 FLPAEVKAKTYIOTELPRKA-KTLHEK-----TSHSSTEACCELCGLYFENRKALASHAR 721  
 QY 131 RSLAHP 136  
 Db 722 AHLROF 727

RESULT 10  
 T00248  
 zinc finger protein wizl - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 05-Nov-1999  
 R:Matsumoto, K.; Ishii, N.; Yoshida, S.; Shiosaka, S.; Watanaka, A.; Tohyama, M.  
 submitted to the EMBL Data Library, March 1998  
 A:Description: Molecular cloning and distinct developmental expression pattern of splice  
 A:Reference number: 214130  
 A:Accession: T00248  
 A:Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1561 <MAT>  
 A:Cross-references: EMBL:AB012265; NID:d1227740; PIDN:BAA32790.1; PID:d1033756  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: wizl

Query Match  
 Best Local Similarity 7.1%; Score 87; DB 2; Length 1561;  
 Matches 35; Conservative 19; Mismatches 48; Indels 24; Gaps 7;

QY 22 LPVPAALNRGTDGPGPSIQKTYDTRYLEHQLRLAG---TYLNYLGP-----PENEPPD 73  
 Db 1221 LPLSPPLASRPKPGAGPT-----QVPR--ELSLSPITSKPSAASYLGVPATKRPLQEDR 1273  
 QY 74 FNPRLGAETLPRAVDLEWMSLNDKLRITONTYEAASHLLCYLRGL---NRQATATELR 130  
 Db 1274 FLPAEVKAKTYIOTELPRKA-KTLHEK-----TSHSSTEACCELCGLYFENRKALASHAR 1327  
 QY 131 RSLAHP 136  
 Db 1328 AHLROF 1333

RESULT 11  
 T34916  
 transferase - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 02-Sep-2000  
 C:Accession: T34916  
 R:Olliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Kaulstrom, M.A.  
 submitted to the EMBL Data Library, January 1998  
 A:Reference number: 221558  
 A:Accession: T34916  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-640 <COLT>  
 A:Cross-references: EMBL:AL021409; PIDN:CAI6181.1; GSPDR:MG 070; SCQEDB:SC3F7.10  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCQEDB:SC3F7.10  
 C:Superfamily: glycine C-acetyltransferase homology  
 F:287-624/Domain: glycine C-acetyltransferase homology <GFA>

Query Match  
 Best Local Similarity 7.1%; Score 86.5; DB 2; Length 1;  
 Matches 50; Conservative 13; Mismatches 64; Indels 57; Gaps 9;

QY 22 LPVPAALNRGTDGPGPSIQKTYDTRYLEHQL---RSLAGTYLNYLGPENPPDPPR 78  
 Db 100 LPVPA-EGTGDPVTEAVAAAROTYORHQLGADLEG-----E 140  
 QY 79 LGAETLPRAVDLEWMSLNDKLRITONTYEAASHLLCYLRGL---NRQATATELR 138  
 Db 141 LGVDSVLTNSVAEA---TERLGLT-----GAAPDAKATTTALAA---D 179  
 QY 139 SLQGLLGSTAGYMAALGYPLPPLPGTEPTWPGPAHSDFLQKMDPFWI1KEIQTWLMRS 198  
 Db 180 ALRGL-----VAAAPGTAPEAPATGAA-APAGRSNAP-----APCAIXGDIHRS 225  
 QY 199 AKDF 202  
 Db 226 MKDF 229

RESULT 12  
 C38604  
 poly(3-hydroxyalkanoate) polymerase (EC 2.7.7.-) 3 - Pseudomonas oleovorans  
 C:Species: Pseudomonas oleovorans  
 C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 08-Sep-2000  
 R:Huisman, G.W.; Monink, E.; Meima, R.; Kazemier, B.; Terpstra, P.; Wilholt, B.  
 J. Biol. Chem. 266, 2191-2198, 1991  
 A:Title: Metabolism of poly(3-hydroxyalkanoates) (PHAs) by Pseudomonas oleovorans. Id  
 A:Reference number: A38604; MUID:91115830  
 A:Accession: C38604  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-560 <HUI>  
 A:Cross-references: GB:M58445; NID:g151441; PIDN:AAA25934.1; ...g151444  
 C:Superfamily: poly(3-hydroxyalkanoic acid) synthase phbC  
 C:Keywords: nucleotidyltransferase

Query Match  
 Best Local Similarity 7.0%; Score 85.5; DB 2; Length 560;  
 Matches 38; Conservative 17; Mismatches 42; Indels 29; Gaps 7;

QY 77 PRGAETLPRAVDLEWMSLNDKLRITONTYEAASHLLCYLRGLNRQATATELRSL--A 134  
 Db 5 PAKGTPLPATSMNVQ-----NAITGLRGR-----DITSTRNVSROS---LRHPLHTA 50  
 QY 135 HFCTSLQGLLGSTAGYMAALGYPLPPLPG-----TEPTWPGPAHSDFLQKMDPFWLKE 190  
 Db 51 HHLALGOLGRV-----ILGDTPLQPNDRPSPDPWQGNPFYKRGLOA---YLAQW 101  
 QY 191 LQTWLM 196  
 Db 102 KOTRLM 107

RESULT 13  
 F72511  
 Probable: cystathionine gamma-synthase APE2068 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: F72511  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999  
 A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
 A:Reference number: A72450; MUID:99310339  
 A:Accession: F72511  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-389 <KAM>  
 A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAAB1078.1; PID:g5105766  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2068  
 C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match	6.9%	Score 84.5;	DB 2;	Length 389;
Best Local Similarity	22.9%	Pred. No. 3.9;		
Matches	47;	Conservative 19;	Mismatches 78;	Indels 61;
				Gaps 8;
QY	43	TYDLTRYLEHQLRSAGTYLNTYLGPPFNPDENPRIGAEPLPRATVDEVMRSI-----	97	
Db	107	TYGSRSLRLEMLSTGTGLEVRLAGPWN-----	LLDLVCADADLIYES	150
QY	98	--NDKRLRTQNEAVSHLLCYLRGLNRQATAELE-----	RSL--AHFCTS	139
Db	151	MANPLRLRPP-----	LSGIRREGSGGVVYVDNFTFATPIAYRPLERGAHY--S	197
QY	140	LOGLLGSLTAGVMAALGYPLPOLP-EGTEPTWPGFAHSDFLQKMDFMILKELQWLMRS		198
Db	198	LESLTRYTAGHNDVYGGSLSGRVEEDLEPLNMNRILCTINQPIDAY-----	LANRG	249
QY	199	AKDENRLKKKQOPPAAAVTLHLGAH	223	
Db	250	MKTLKARFENQSRRAVEVAEWLESH	274	

RESULT 14  
F75439  
probable cell wall synthesis protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: F75439  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.D.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Ulterback, T.; Zalewski, C.; M.  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: F75439  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-411 <WHI>  
A:Cross-references: GB:AE001958; GB:AE000513; NID:g6458805; PIDN:AF10649.1; PID:g645880  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1076  
A:Map position: 1

```

Query Match      6 9%; Score 84; DB 2; Length 411;
Best Local Similarity 24.68; Pred. No. 4.6;
Matches 56; Conservative 25; Mismatches 77; Indels 70; Gaps 11;

OY 2 DLRAGDSMGLACTVLMHLPAVPALNRCTDPPGPSIKQTDLTFTYLEHQRLSLAGTY 61
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Db      212 EVLADRMALRALPSEMSLRGPVLLLLSG -GRGHYAAAADVL*-----fLGNL- GRA 264
QY      62 LNYLGPPEPPNEPPNPRLGAEITLPRAIVDLEWMSLNDKRLTQNYEASH*CYLKGIN 121
Db      265 VQVLPAPASRQGE-----GTETIGATV-----Hh*FFRRDLR 236
QY      122 RQAAVAELRSLAHFCTSLQSLGINSV---MAALGYLP-QPLP*H* *WGPAP 175
Db      297 RLAAASDL-----VKKAGGLVAEATATGAPLVYAPAPQGE-----EHN 337
QY      176 SDFLOKMDPEWLKLELOTWIM-NSAKDENRLKKMKQPPAAAVTLHGCA 222
Db      338 ADPLERHG-----AGLWARRHHDVRLVLRALDPRDHAHLASACA 376

```

RESULT 15  
71327  
probable apolipoprotein N-acyltransferase (cute) - syphilis spirochete  
C:Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 17-Mar-2000  
C:Accession: G71327  
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutcliffe, G.G.; Dodson, R.; Gerson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; W  
they, L.; Weisman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A:Reference number: A71250; M01D:9832770  
A:Accession: G71327  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-559 <COL>  
A:Cross-references: GB:AE001219; GB:AE000520; NID:g3322693; P:DN:AAAC65404.1; PID:g3332  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0417  
C:Superfamily: apolipoprotein N-acyltransferase

Query Match	6.9%	Score 84	DB 2	Length 559
Best Local Similarity	21.6%	Pred. No. 6.8		
Matches	48	Conservative	17	Mismatches 75; Indels 82; Gaps 9;

QY	10	GMLACLC-----	TVLMLPAVALNLRKCPGSPISOKRYDITRL	EHQ	53
Db	197	GCYLCLSLYGAVRIATLGCAPRTL	---ALAIYQNMADPMTWTSFERNLTTHAIHLTEA		252
QY	54	LRSLAGTYLNIYGPENPEDENPRLGCAETLPRATVLDLFWBSLNDKLRLLQNYEAVSHL			113
Db	253	LRT-----QTAPRLPTTPYKRKEKTLTHASARAFAVDVMV	---SESSLRPYQYRHHV		301
QY	114	L-----CYLGLNLRQA-----	ATAELRRLSLAHFCTSLQGL		144
Db	302	YNALPAPRPFASFLRTLGAPLLVGTPLRLSGNSNTGGVANNVALLRP	-----DGHV		352
QY	145	GSIAGVMALGXPR-----	LPQPLPGTEPIWTPGP		173
Db	353	AOYTGKQIMVPAEFLPWGHTSVQURLQIMLAGSESSTWPGP			394

Search completed: March 18, 2002, 09:32:30  
Job time: 216 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:31:24 ; Search time 29.41 Seconds

(Without alignments)  
280.503 Million cell updates/sec

Title: US-09-931-704-2

Perfect score: 1226

Sequence: 1 MDLRAGDSMGLACTVLM.....KKKQPPAAAVTLHGANGF 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	124.5	10.2	203 1	CTFL_RAT
2	118.5	9.7	203 1	CTFL_MOUSE
3	91.5	7.5	201 1	CTFL_HUMAN
4	91.5	7.5	1182 1	HAIR_MOUSE
5	91	7.4	195 1	CNTE_CHICK
6	89.5	7.3	1313 1	MIPL_SCHPO
7	87.5	7.1	1181 1	HAIR_RAT
8	85.5	7.0	560 1	PHAC_PSEOL
9	84	6.9	200 1	CNTE_PIG
10	83	6.8	452 1	TLL_DROME
11	82.5	6.7	1189 1	HAIR_HUMAN
12	82	6.7	724 1	P85B_BOVIN
13	81	6.6	291 1	TM32_MYCTU
14	80.5	6.6	294 1	HOPK_RHOCA
15	80.5	6.6	870 1	BCAL_HUMAN
16	80	6.5	200 1	CNTE_RAT
17	80	6.5	422 1	Y140_HUMAN
18	80	6.5	2261 1	RRPL_MOUSE
19	78.5	6.4	390 1	YGLI_STRKO
20	77	6.3	586 1	UL84_HCVTA
21	76.5	6.3	587 1	UL84_HCVTA
22	76.5	6.2	571 1	ATKA_MYCTU
23	76.5	6.2	572 1	SYM_AERPE
24	76.5	6.2	1006 1	MANE_PIG
25	76	6.2	1001 1	PTPX_MOUSE
26	76	6.2	1132 1	BAT3_HUMAN
27	75.5	6.2	830 1	VPP3_HUMAN
28	75	6.1	199 1	CNTE_RABIT
29	75	6.1	343 1	DFRA_SYNT3
30	75	6.1	450 1	TLL_DROVI
31	75	6.1	917 1	STL_STRAV
32	74.5	6.1	346 1	YG2Q_YEAST
33	74.5	6.1	444 1	CEFG_CEPAC

34	74	6.0	508 1	EGR1_RAT	54	rattus norv
35	74	6.0	859 1	YDBB_SCHPO	51	462 schizosacch
36	73.5	6.0	829 1	DXS_AQUAE	067046	aquifex aeo
37	73.5	6.0	1009 1	PAK2_HUMAN	014289	homo sapien
38	73	6.0	372 1	CD14_RABIT	028780	homo sapien
39	73	6.0	39	DUS9_HUMAN	097956	homo sapien
40	73	6.0	384 1	YNTB_ECOLI	P7 323	escherichia
41	73	6.0	531 1	CP27_HUMAN	067118	h cytochrom
42	73	6.0	627 1	DXS_PSEAE	097107	pseudomonas
43	73	6.0	874 1	BCAL_MOUSE	061140	mus musculu
44	73	6.0	968 1	BCAL_RAT	063767	rattus norv
45	72.5	5.9	276 1	END8_STRKO	086820	streptomyces

## ALIGNMENTS

RESULT	1	STANDARD:	PRT:	203 AA.
CTFL_RAT				
ID	CTFL_RAT			
AC	063086;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	CARDIOTROPIN-1 (CT-1).			
GN	CTFL			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-MISTAR: TISSUE-Heart;			
RX	MEDLINE=6613659; PubMed=8604995;			
RA	Ishikawa M., Saito Y., Miyamoto Y., Kuwahara K., Ogawa T.			
RA	Nakagawa O., Harada M., Masuda I., Nakao K.;			
RT	"CDNA cloning of rat cardiostrophin-1 (CT-1): augmented expression of			
RT	CT-1 gene in ventricle of genetically hypertensive rats."			
RL	Biochem. Biophys. Res. Commun. 219:377-381(1996).			
CC	-1- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO			
CC	AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF			
CC	RECEPTOR)/GP 130 RECEPTOR COMPLEX.			
CC	-1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRICLE AND MUSCLE OF ADULT			
CC	RATS. ALSO DETECTED IN THE LUNG, KIDNEY, LIVER, SKELETAL MUSCLE,			
CC	STOMACH AND URINARY BLADDER. NOT DETECTED IN BRAIN, THYMUS, TESTIS,			
CC	SPLEEN OR THYMUS. OVEREXPRESSED IN THE VENTRICLES IN THE CASE OF			
CC	HYPERTENSION AND HYPERTROPHY.			
CC	-1- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isi.stb.ch/announce/">http://www.isi.stb.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@stb.ch">license@stb.ch</a> ).			
CC	-----			
DR	EMBL; D78591; BA11427.1; .			
DR	InterPro; IPR001581; LIF_OSM.			
DR	SMART; SM00080; LIF_OSM; 1.			
KW	Cytokine.			
SQ	SEQUENCE 203 AA; 21439 MW; DF8B921A2FA1C832 CRC64;			
Query Match	10.2%; Score 124.5; DB 1; Length 203;			
Best Local Similarity	28.1%; Pred. No. 0.00011;			
Matches	50; Conservative 24; Mismatches 85; Indel. 10; Gaps 5;			
QY	40 IOKYDITRLIEHLRLSLAGTYLNYLGPPEPDENPRP---GAETLPRAIVDLRWMS 96			
DB	27 IROTHNLRLITKYADQLLEEVYQOQGEPPGLGFSPPRLAGLSGAPARSINAGLPV--- 83			

OY	97	LNDKRLTDPNVEAFSHLLCYLGLUNRQAA-----TAELRRSLAHFCTSILOGLIGSIAGYM	151
Dd	84	---SERLRDQAALSALEPALLDAVVRROAEINPRAPRLRLSRLEDARQRALGAAYETVL	140
OY	152	AALGY----PLDPGLETEPTWTPGPAHSDFLQKMDDFWLKELOTWLMRSKADFNRL	205
Dd	141	AALGAARBPVEBVP-ATSALEFTSNASAGVFAKYLGIVHCILYGEMWSRTGEDIGQL	197
<hr/>			
RESULT 2			
ID	CTFL_MOUSE	STANDARD:	PRT: 203 AA.
AC	Q60753;		
DR	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DI	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	CARDIOTROPHEIN-1 (Ct-1).		
GN	CTFL.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9516785; PubMed=7862649;		
RA	Pennica D., King K.L., Shaw K.J., Luis E., Rullamas J., Luoh S.-M.,		
RA	Darbonne W.C., Knutson D.S., Yen R., Chien K.R., Baker J.B.,		
RD	Wood W.I.;		
RT	"Expression cloning of cardiotrophin 1, a cytokine that induces		
RL	cardiac myocyte hypertrophy."		
RU	Proc. Natl. Acad. Sci. U.S.A. 92:1142-1146(1995).		
CC	-1- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO		
CC	AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF		
CC	RECEPTOR)/GP 130 RECEPTOR COMPLEX.		
CC	-1- SUBCELLULAR LOCATION: SECRETED.		
CC	-1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,		
CC	LIVER, LUNG AND KIDNEY. LOWER LEVELS IN TESTIS AND BRAIN. NO		
CC	EXPRESSION IN SPLEEN.		
CC	-1- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; U18366; AAC52173.1; .		
DR	MGI:105115; Ctf1.		
KW	Cytokine.		
SQ	SEQUENCE 203 AA: 21509 MW: 883DA14A0B38232F CRC64:		
<hr/>			
	Query Match	9.7%; Score 118.5; DB 1; Length 203;	
	Best Local Similarity	28.8%; Pred. No. 0.00041;	
	Matches 51; Conservative 21; Mismatches 88; Indels 17; Gaps 5;		
OY	40	IOKTIDLTRYLEHOURLSLAGTYLANTLPPEPNPDNPRL---GAEYLPRATVDLEWRMS	96
Dd	27	IROTHTNLARLLTKVAEQOLLEEVOOGEFFPGSPRLPLAGLSGAPASHAGIPV--	83
OY	97	LNDKRLTDPNVEAFSHLLCYLGLRROAA-----TAELRRSLAHFCTSILOGLSIAGYM	151
Dd	84	---SERLRDQAALSALEPALLDAVVRROAEINPRAPRLRLSRLEDARQRALGAAYETVL	140
OY	152	AALGYPL---POPUPCTEPT-WTPGPAHSDFLQKMDDFWLKELOTWLMRSKADFNRL	205
Dd	141	AALGAARBPVEBVPVATLTFTANSTAGIFSAXVIGFHVCILYGEMWSRTGEDIGQL	197
<hr/>			
RESULT 3			
CTFL_HUMAN			

ID	CTFL HUMAN	STANDARD:	PRT:	201 AA.
AC	016619;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	CARDIOTROPIN-1 (CT-1).			
GN	CTFL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homu.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Heart;			
RX	MEDLINE=96429882; PubMed=8833032;			
RA	Penhica D., Skanson T.A., Shaw K.J., Kuang W.-J., Gray C.L.,			
RA	Beatty B.G., Wood W.I.;			
RT	"Human cardiortrophin-1: protein and gene structure, biological and			
RT	binding activities, and chromosomal localization.";			
RL	Cytokine 8:183-189(1996).			
CC	-I- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERPROLIF IN VITRO. BINDS TO			
CC	AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF			
CC	RECEPTOR)/GP 130 RECEPTOR COMPLEX.			
CC	-I- SUBCELLULAR LOCATION: SECRETED.			
CC	-I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,			
CC	PROSTATE AND OVARY. LOWER LEVELS IN LUNG, KIDNEY, PANCREAS,			
CC	THYMUS, TESTIS AND SMALL INTESTINE. LITTLE OR NO EXPRESSION IN			
CC	BRAIN, PLACENTA, LIVER, SPLEEN, COLON OR PERIPHERAL BLOOD			
CC	LEUKOCYTES.			
CC	-I- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U43033; AAD12173.1;			
DR	EMBL; U43031; AAD12173.1; JOINED.			
DR	EMBL; U43032; AAD12173.1; JOINED.			
DR	EMBL; U43030; AAB85229.1;			
DR	HSSP; P04002; IMFA.			
DR	MIM; 600435; -.			
KW	Cytokine.			
SO	SEQUENCE 201 AA; 21227 MW; 0235A7B5745F675F CRC64;			
Query Match: 7.5%; Score 91.5; DB 1; Length 401;				
Best Local Similarity 25.9%; Pred. No. 0.14;				
Matches 45; Conservative 24; Mismatches 92; Indels 13; Gaps 5;				
QY	40 IOKTYDILRYIEHQRLSLAGTLYLNGPFNPDPNPRPLGAETLPPRAIVDLVWRSIND 99			:::
DB	27 IQHTSHLHLKRYKAEQQLQEVYQGDGFGLSPSPRLPYAGL-SAPAPSHADLPVHE 85			:::
QY	100 KIRLTONTYAVSHLLCYLRGLNRQA-----TAELRSLAHFCTSLQGLGSIAGVMAAL 154			:::
DB	86 RLRLL--DAAALALAPLLLDAYACRQAEINPRAPRLRLRLLEDAARQARAI;GAAVALLAAL 143			:::
QY	155 GYVPLPOLPLPGTPE--TWTPGPAHSDPLQKMDDEFLKELQTVLWRSKKDFNRL 205			:::
DB	144 G--AANRGPRAEPPEPATYASASATGYFFAKYVGLGRVCGILREWLSTRIGDILQQL 195			:::
RESULT 4				
HAIR_MOUSE				
ID	HAIR_MOUSE	STANDARD:	PRT:	1182 AA.
AC	061645;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			

```

DE HAIRLESS PROTEIN.
OS HR.
OS Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=Skin;
RX MEDLINE=94329587; Pubmed=8052649;
RA Cachon-Gonzalez M.B., Fennel S., Coffin J.M., Moran C., Best S.,
RT "Structure and expression of the hairless gene of mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).
CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO
CC -1- REGULATE ONE OF THE PHASES OF HAIR GROWTH.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN, HAIR
CC FOLLICLES AND INTERFOLLICULAR EPIDERMIS. NO EXPRESSION IN
CC DERMIS.
CC -1- DISEASE: HR MUTATION PRODUCES A NUMBER OF PLEIOTROPIC EFFECTS
CC INCLUDING STRUCTURAL ABNORMALITIES OF EPITHELIAL CELLS IN THE HAIR
CC FOLLICLES, HAIR LOSS TOWARDS THE END OF THE FIRST HAIR GROWTH
CC CYCLE, AND THE FAILURE OF SUBSEQUENT HAIR GROWTH CYCLES. OLDER
CC MICE CARRYING AN HR MUTATION HAVE BEEN REPORTED TO POSSESS ALTERED
CC RATIOS OF T-CELL-DEPENDENT B-CELL RESPONSES. MICE HOMOLOGOUS FOR
CC HR MUTATION ARE UNIQUELY SENSITIVE TO UV AND CHEMICALLY INDUCED
CC SKIN TUMORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: Z32675; CAA83587.1; -.
DR MGD: MGI:96223; hr.
DR Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
KW Metal-binding.
KW ZN_FING. 595 620 C6-TYPE.
FT DOMAIN 535 540 POLY-GLY.
FT SEQUENCE 1182 AA; 127182 MW; EFBB0EE62AE81B40 CRC64;
SQ
Query Match 7.5%; Score 91.5; DB 1; Length 1182;
Best Local Similarity 23.7%; Pred. No. 1.3;
Matches 70; Conservative 25; Mismatches 99; Indels 101; Gaps 17
3 LRAGDSWGLACLTVMHRAVVALNRTPRG-----PGP----- 38
1111 1111 1111 1111
DB 813 LRAGS--GLRKQL-----SLRLSPRTSLSPGALLMLQERPRKNGHLEQEHNRQGRV 865
OY 3 LRAGDSWGLACLTVMHRAVVALNRTPRG-----PGP----- 38
1111 1111 1111 1111
DB 866 LVSGIOKTRLRLSMGMEMALGLGGVQVLTALGPQPNLDSTAFWEGFSHPERP-- 921
OY 39 ---STCKTYDLRYLEHQRSLAGTY--LNYLGRP-----ENEDPFPRLG 80
1111 1111 1111 1111
DB 81 AETPLRAYVDLEWMSLNDK-LRLTON-----YEAYSHLCYL-RGLNRQAAAT 126
OY 81 AETPLRAYVDLEWMSLNDK-LRLTON-----YEAYSHLCYL-RGLNRQAAAT 126
1111 1111 1111 1111
DB 922 --KLDGGSV-LLIHRTLGRKASRYONLASSLPDPEYCAHOGKLNLSYDLGLTLPLE 978
OY 922 --KLDGGSV-LLIHRTLGRKASRYONLASSLPDPEYCAHOGKLNLSYDLGLTLPLE 978
1111 1111 1111 1111
DB 127 AEL-----RRSLAHFCTSLQGLGSIAGVMAALGYRLPOLRDLPTERTPTPGRAH 175
OY 127 AEL-----RRSLAHFCTSLQGLGSIAGVMAALGYRLPOLRDLPTERTPTPGRAH 175
1111 1111 1111 1111
DB 979 POLMAAYGVNHRGHGTLGTNLCEVSDLSIVHAEOQ-----PPWY--RAQ 1024
OY 979 POLMAAYGVNHRGHGTLGTNLCEVSDLSIVHAEOQ-----PPWY--RAQ 1024
1111 1111 1111 1111
DB 176 SDFLQKMD--DFULLKELOTLMR--SADPFRILKKMP--PAAVTLHLGANG 224
OY 176 SDFLQKMD--DFULLKELOTLMR--SADPFRILKKMP--PAAVTLHLGANG 224
1111 1111 1111 1111
DB 1025 KDFLSGLDGGGLTSPSGQSTVWVHRPADQAIRKRFLOMCPDAGAGLTLEPGAPG 1079
OY 1025 KDFLSGLDGGGLTSPSGQSTVWVHRPADQAIRKRFLOMCPDAGAGLTLEPGAPG 1079

```

```

ID CNMF.CHICK STANDARD; PRT; 195 AA.
DT 002011:
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1996 (Rel. 34, Last annotation update)
DE CILIARY NEUROTROPHIC FACTOR (CNTF) (GROWTH PROMOTING ACTIVITY) (GPA).
GN CNTF OR GPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. PubMed-1610564.
RX Leung D.W., Parent A.S., Cachianes G., Lee A.L., Nikolic K.,
RA Esch F., Coulombe J.N., Blacher R.W., Eckenstein F.P., Nishii R.;
RT "Cloning, expression during development, and evidence for release of
RT a trophic factor for ciliary ganglion neurons."
RL Neuron 8:1045-1053(1992).
CC -1- FUNCTION: CNTF IS A SURVIVAL FACTOR FOR VARIOUS NEURONAL CELL
CC TYPES. SEEMS TO PREVENT THE DEGENERATION OF MOTOR AXONS AFTER
CC AXOTOMY.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE CNTF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isdb.ch/announce/
CC or send an email to license@isdb.ch)).
CC -----
CC EMBL: M80827; AAA48784.1; .
CC PIR: JH0680; JH0680.
DR HSPSP_P26441; ICNT.
DR InterPro: IPR000151; CNTF.
DR Pfam: PF01110; CNTF_1.
DR ProDom: PD011041; CNTF_1.
KW Growth factor; Neurone.
SQ SEQUENCE 195 AA; 21330 MW; FEAD76949DB34AC5 CRC64.

Query Match 7.4%; Score 91; DB 1; Length 195;
Best Local Similarity 27.3%; Pred. No. 0.15;
Matches 51; Conservative 21; Mismatches 83; Indels 32; Gaps 9

46 LRRYLEHQLRSIAGTYLNLGLPPRFNEPFPNPRRGAETLPATVDLEWV To KLRIHQ 105
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
23 LARKMSDVTDLDDLYVERQ-----LDASISVAANDVGPTAAV- ERMAVLITFORILD 75
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 N---YEAVSHLCYLGRINLRQA--ATAELRRSLA-----HECTSLGGCI SIACVMA 152
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
76 NLAAAYAFETLLAQMLIEQRRELGTDDELGPALAMLLQVSAPFYHHLEL----ELE 130
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
153 ALGYLPQPLPGTEPWTPGRAH-SDFLOKMDDEWLKELOFTWLRSAKDI NKLKKMP 211
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 SRGAPAE--SESP--PAPRPLSFEOKLGIVLKEFLMGVMANVRSDRLDPLSKHGPG 184
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
212 PAAYTL 218
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 SGAALGL 191
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
MIP1_SCHPO STANDARD; PRT; 1313 AA.
ID MIP1_SCHPO
AC P87141;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
```

DE WD-REPEAT PROTEIN MIP1.  
 GN MIP1 OR SPAC57A7.11.  
 OS Schizosaccharomyces pombe (fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 NCBI\_TaxId=4896;

RP [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=20115869; PubMed=10648609.  
 RA Shinozaki-Yabana S., Matanabe Y., Yamamoto M.;  
 RT "Novel WD-repeat protein Mip1 facilitates function of the meiotic  
 RT regulator Meip1 in fission yeast."  
 RL Mol. Cell. Biol. 20:1234-1242(2000).

RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Skelton J., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;  
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: BINDS TO AND FACILITATES THE FUNCTIONING OF THE MEIOTIC  
 CC REGULATOR MEI2. MAY ALSO BE INVOLVED IN CONJUGATION BY INTERACTING  
 CC WITH STE11. ESSENTIAL FOR CELL GROWTH.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB032552; BAA84585.1;  
 DR EMBL; Z95396; CAB08769.1;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR SMART: SM00320; WD40; 5.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_REGION; 1.  
 KW Meiosis; WD repeat; Repeat.  
 FT REPEAT 986 1029 WD 1.  
 FT REPEAT 1033 1074 WD 2.  
 FT REPEAT 1087 1126 WD 3.  
 FT REPEAT 1130 1170 WD 4.  
 FT REPEAT 1176 1216 WD 5.  
 FT REPEAT 1219 1259 WD 6.  
 FT REPEAT 1268 1308 WD 7.  
 SQ SEQUENCE 1313 AA; 148533 MW; C71B663B0171E7A4 CRC64;

Query Match 7.3%; Score 89.5; DB 1; Length 1313;  
 Best Local Similarity 22.8%; Pred. No. 2.4;  
 Matches 61; Conservative 32; Mismatches 69; Indels 105; Gaps 14;

QY 15 LCTVLMHLPAAV--PALNR-----TGDPGPPST-----OKTYDLTRYLEHQLRS 56  
 DB 811 LAFLLQHPALHKKSLSKDTPNSVTSQPKPHFPVPSYSENKILNRSFSLRSLKGLALS 870  
 QY 57 LAG-----TYLNYL-----GPP--FNEPDEFNPRILGATLPRAVYL 91  
 DB 871 LAGSDRASSELLNGENKPAESNHLHLSAKVPPAPAFNELEY-----QSELD 919  
 QY 92 EWMRLND-----KLRLTQYEAYSLLCYLRGL-----NRQATVAELRSLAHFCTS 139  
 DB 920 PLTSTLFDWMSKKYTERPQMRNEDDEPGS--ICYNRLMRNRNEKLIYTRPLAEVSTN 977  
 QY 140 -----DGLGLGSIAGYMAALGYPLPQLPPTPTWTPGPAHSDFLQKMDFWLKE--- 190  
 DB 978 GRNQQLMTFNNTIA-----PRKLMTFQFEDQLITLGDKDI 1013

QY 191 LOTWLRSAKDFNRLKKKMQPPAAVT 217  
 DB 1014 IQVWDMRR-----NRLNSEKTSASATV 1036

RESULT 7  
 HAIR\_RAT  
 ID HAIR\_RAT STANDARD; PRT; 1181 AA.  
 AC P97609;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HAIRLESS PROTEIN.  
 GN HR.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxId=10116;

RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SEPRAGUE-DAWLEY;  
 RX MEDLINE=97141510; PubMed=8987811;  
 RA Thompson C.C.;  
 RT "Thyroid hormone-responsive genes in developing cerebellum include a  
 RT novel synaptotagmin and a hairless homolog."  
 RL J. Neurosci. 16:7832-7840(1996).  
 CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COU  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U71293; AAC53018.1; ALT\_INIT.  
 KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;  
 KW Metal-binding;  
 FT ZN\_FING 594 619 C6-TYPE.  
 SQ SEQUENCE 1181 AA; 127307 MW; 834B7029CF8E88F0 CRC64;

Query Match 7.1%; Score 87.5; DB 1; Length 1,81;  
 Best Local Similarity 25.3%; Pred. No. 3.2;  
 Matches 61; Conservative 21; Mismatches 92; Indels 67; Gaps 12;

QY 30 RTGPGGPPSTIQKTYDLTRYLEHQLRSIAGTY--LNYLGP-----FNEPDP 74  
 DB 859 RQGPPVLVSGIQKTLRLSLWGMEMALGTLGGQVQTLTALGPPQPSLSDSVAFKGFSPPEA 918  
 QY 75 NPPRLGAEETPRAVYDELVNRSINDK-----LRLTO-----NEAVSHLLC 115  
 DB 919 RP-----KIDEGSV--LLHRLPLGDKDESREVENIASSLPLEVCAGHQKLNILASVPLDL 971  
 QY 116 YLNGLNQAAATLAELRS-----LAHFTSLQGLGSIAGYMAALGYPLPQLPPTPTW 169  
 DB 972 TLHPLEQLMAAGYVNSHSHGLSTKNLCVEVSDLLILVHAERQL-----PPW 1019  
 QY 170 TPGPAHSDFLQKMD--DFWLKELQTLWLR--SAKDNLKRRMP--PAAAVTLHLGHN 223  
 DB 1020 Y--RAQKDFLSIGDGLGSLWSPQSOTSTVHVPRAQDAQIRRFLLQVCPAGAGTLEPGAP 1077  
 QY 224 G 224  
 DB 1078 G 1078  
 RESULT 8  
 PHAC\_PSEOL STANDARD; PRT; 560 AA.  
 ID PHAC\_PSEOL



AC P26496;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE POLY(3-HYDROXYALKANOATE) POLYMERASE 2 (EC 2.3.1.-) (PHA POLYMERASE 2)  
 DE (PHA SYNTHASE 2) (POLYHYDROXYALKANOIC ACID SYNTHASE 2).  
 GN PHAC.  
 OS Pseudomonas oleovorans.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 OX NCBI\_Taxid=301;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GP01;  
 RX MEDLINE=91115830; PubMed=1969978;  
 RA Huisman G.W., Wouink E., Melma R., Kazemier B., Terpstra P.,  
 RA Wiltbold B.;  
 RT "Metabolism of poly(3-hydroxyalkanoates) (PHAs) by Pseudomonas  
 RT oleovorans. Identification and sequences of genes and function of the  
 RT encoded proteins in the synthesis and degradation of PHA.";  
 RL J. Biol. Chem. 266:2191-2198(1991).  
 CC -1- FUNCTION: POLYMERASE ACCUMULATES POLY(3-HYDROXYALKANOATES) AFTER  
 CC GROWTH ON MEDIUM CHAIN LENGTH HYDROCARBONS. LARGE AMOUNTS OF THIS  
 CC POLYESTER ARE SYNTHESIZED WHEN CELLS ARE GROWN UNDER NITROGEN-  
 CC LIMITING CONDITIONS. WHEN NITROGEN IS RESUPPLIED IN THE MEDIUM,  
 CC THE ACCUMULATED PHA IS DEGRADED.  
 CC -1- SIMILARITY: BELONGS TO THE PHA/PHB SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 CC EMBL: M58445; AAA25934.1; -  
 DR PIR: C38604; C38604.  
 DR InterPro: IPR000073; Abhydrolase.  
 DR Pfam: PF00561; abhydrolase; 1.  
 KM PHA biosynthesis; Transferase; Acyltransferase.  
 FT ACT\_SITE 296 POTENTIAL.  
 SQ SEQUENCE 560 AA; 62631 MW; E2CD844FC1616B83 CRC64;

Query Match 7.0%; Score 85.5; DB 1; Length 560;  
 Best Local Similarity 30.2%; Pred. No. 1.9;  
 Matches 38; Conservative 17; Mismatches 42; Indels 29; Gaps 7;

QY 77 PRGAEPLPATYDLEWMSLNDKLTQNYEAYSHLCLYGLNRQAATAEIRSL--A 134  
 Db 5 PAKGTPLPTSMNVQ-----NAILGLRGR-----DLISLKNVSRQS-----LRHPDLTA 50  
 QY 135 HFCTSLGGLGSIAGVMAALGYPLPQPLPG---TEPTWPGPAHSDPLQKKMDFWLKE 190  
 Db 51 HHLALGGLGGRV-----ILGDTPLQPNRPDPSPFTWSQNFYRGLQA---YLAWO 101  
 QY 191 LQTWLM 196  
 Db 102 KQTRLM 107

RESULT 9  
 CNTF\_PIG STANDARD; PRT; 200 AA.  
 ID CNTF\_PIG  
 AC 002732;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CILIARY NEUROTROPHIC FACTOR (CNTF).  
 GN CNTF.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus  
 OX NCBI\_Taxid=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98030048; PubMed=9363597;  
 RA Weaks R.L., Ramsondar J.J., Gallagher D.S. Jr., Nogueira L.,  
 RA Piedrahita J.A.;  
 RT "Isolation, characterization and chromosomal localization of the  
 RT porcine ciliary neurotrophic factor (CNF) gene.";  
 RL Anim. Genet. 28:354-357(1997).  
 CC -1- FUNCTION: CNF IS A SURVIVAL FACTOR FOR VARIOUS NEURONAL CELL  
 CC TYPES. SEEMS TO PREVENT THE DEGENERATION OF MOTOR AXONS AFTER  
 CC AXOTOMY (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CNF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 CC EMBL: U57644; AAC27342.1; -  
 DR HSSP: P26441; 1CNT.  
 DR InterPro: IPR000151; CNTF.  
 DR Pfam: PF01110; CNTF; 1.  
 DR ProDom: PD011041; CNTF; 1.  
 KW Growth factor; Neurope.  
 SQ SEQUENCE 200 AA; 22718 MW; 40507C4457ED6531 CRC64;

Query Match 6.9%; Score 84; DB 1; Length 200;  
 Best Local Similarity 24.7%; Pred. No. 0.72;  
 Matches 47; Conservative 23; Mismatches 88; Indels 32; Gaps 9;

QY 46 LREYLEHQLSLAGTYLNYLSPNEPDPNPRGAEPLPATYDLEWMSLNDKLTQ 105  
 Db 23 LARKIRSDELTAIMEAVKHKQ--LINE---NINDSDVDPMASTDR--WSETTAERKLOE 75  
 QY 106 NREAYSHLCLYGLNRQAATAEIRSLAHFCTS---LQGLGSIAGVMAALGYPLPQ-- 160  
 Db 76 NLRAYRFFHYMARL-----LEDQREHFTPADDDHQAHLHIVLVAAFAVQLEELM 127  
 QY 161 -----PLPGTEPTWPGPAHSD--FLQKKMDFWLKEQTLWMSAKQPNRLKKMKQPPA 213  
 Db 128 VLEHKKVPSEAGGTPLSVGGGLFEKKIKMGKLVQLELSQMTVRSIRDIYVSSHQ----- 183  
 QY 214 AAVTLHGAH 223  
 Db 184 AGVPAH-GSH 192

RESULT 10  
 TLL\_DROME STANDARD; PRT; 452 AA.  
 ID TLL\_DROME  
 AC P18102; Q9VA33;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE TAILLESS PROTEIN.  
 GN TLL OR NR2E2 OR CG1378.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Terebrata; Hexapoda; Insecta;  
 OC Eukaryota; Metazoa; Arthropoda; Terebrata; Hexapoda; Insecta;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OREGON-R;  
 RX MEDLINE=90304905; PubMed=2364433;

THE REGULATION OF EARLY EYE DEVELOPMENT. IN THE EMBRYONIC VISUAL SYSTEM ANLAGE DRIVES CELLS TO OPTIC LOBE AS OPPOSED TO BOLWIG'S ORGAN FATE.

```
CC CC -1 SUBUNIT: MONOMER (BY SIMILARITY).
CC CC -1 SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC CC -1 TISSUE SPECIFICITY: BRAIN AND PERIPHERAL NERVOUS SYSTEM.
CC CC -1 DEVELOPMENTAL STAGE: DURING STAGE 10 FOUND IN 1H; ANTERIOR PART OF
CC CC THE VISUAL SYSTEM THAT LATER GIVES RISE TO THE ANTERIOR TIP OF THE
CC CC OPTIC LOBE. AT STAGE 12 ALSO FOUND IN THE POSTERIOR TIP OF THE
CC CC OPTIC LOBE. IN THIRD LARVAL INSTAR EXPRESSED IN THE OPHTHALMIC LOBE OF
CC CC THE LARVAL BRAIN AND IN THE EYE ANTENNAL DISK, BOTH IN ANTENNAL
CC CC AND EYE PORTION.
CC CC -1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC CC NR2 SUBFAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL: M34639; AAA8936.1; -
DR DR EMBL: AF019362; AAB71871.1; -
DR DR EMBL: AE003773; AAP57091.1; -
DR DR PIR: A35602; A35602.
DR DR PIR: A47265; A47265.
DR DR HSSP: P03372; IHCO.
DR DR TRANSFAC: T00789; -.
DR DR FLYBASE: FBgn0003720; t11.
DR DR InterPro: IPR00536; Hormone_rec_1lg.
DR DR InterPro: IPR01628; zf-C4.
DR DR Pfam: PF00104; hormone_rec_1.
DR DR Pfam: PF00105; zf-C4; 1.
DR DR PRINTS: PRO0047; STROIDINGER.
DR DR SMART: SM00430; HOLI; 1.
DR DR SMART: SM00399; ZNF_C4; 1.
DR DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Activator; Repressor; Developmental protein
FT FT DNA_BIND 34 101 NUCLEAR RECEPTOR-TYPE.
FT FT ZN_FING 34 54 C4-TYPE.
FT FT ZN_FING 70 96 C4-TYPE.
FT FT DOMAIN 244 389 LIGAND-BINDING (BY SIMILARITY).
FT FT DOMAIN 262 265 POLY-LEU.
SQ SEQUENCE 452 AA; 50549 MW; A4ABEFFDE993A37C CRC64;
Query Match 6.8%; Score 83; DB 1; Length 452.
Best Local Similarity 23.8%; Pred. No. 2.5;
Matches 34; Conservative 11; Mismatches 32; Indels 66; Gaps 6;
QY 115 CYLRRLNQAAATAE-----LRSLAHFTCTSGLLGSIAGVA-----AALG 155
DB 96 CREVGMNDADVAQHGRPNSTLRHMAMTXKDAWNG-----AGEMPQIPADILMTNTAALG 150
QY 156 YP-LRPGYRTGETPTTPRRASDFLQAKMDFFWLKEQLQTWLMNSAADFNRLKKMPRA 214
DB 151 FRGVGMRYRGRLQRGAAGNHRAHMAFG-----PPSA 181
QY 215 AVTL-----HGANGF 225
DB 182 AAVLDLSVRVRRHNHRVHQGNHGF 204
RESULT 11
HAIR_HUMAN
ID HAIR_HUMAN STANDARD; PRT; 1189 AA.
AC O43593;
DT 15-JUL-1999 (Rel. 38, Created)
DT 20-AUG-1999 (Rel. 38, Last sequence update)
DE HAIRLESS PROTEIN.
OS Homo sapiens (Human).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RX SEQUENCE FROM N.A., AND VARIANT ALONG ALA-1022.  
 RX MEDLINE=98111413; PubMed=9445480;  
 RA Ahmad M., Ul Haque M.F., Brancolini V., Tsou H.C., Ul Haque S.,  
 Lam H., Alta V.M., Owen J., Deblaguerre M., Frank J.,  
 Raeserhaml-Friedman P.B., Leask A., McGrath J.A., Peacocke M.,  
 RA Ahmad M., Ott J., Christiano A.M.;  
 RT "Alopecia universalis associated with a mutation in the human hairless  
 gene";  
 RT Science 279:720-724(1998).  
 RL [2]  
 RN VARIANT APL GLN-620.  
 RP MEDLINE=98431781; PubMed=9758627;  
 RX Ahmad M., Irvine A.D., Lam H., Buckley C., Bingham E.A.,  
 RA Partridge A.A., Ahmad M., McGrath J.A., Christiano A.M.;  
 RT "A missense mutation in the zinc-finger domain of the human hairless  
 gene underlies congenital atrichia in a family of Irish travellers";  
 RT Am. J. Hum. Genet. 63:984-991(1998).  
 RL [3]  
 RN VARIANT ALONG ASP-1136, AND ALTERNATIVE SPLICING.  
 RP MEDLINE=98409496; PubMed=9736769;  
 RX Cichon S., Anker M., Vogt I.R., Kohleder H., Putzstuck M., Hillmer A.,  
 RA Farooq S.A., Al-Dhafiri K.S., Ahmad M., Haque S., Riettschel M.,  
 RA Propping P., Kruse R., Nothen M.M.;  
 RT "Cloning, genomic organization, alternative transcripts and mutational  
 analysis of the gene responsible for autosomal recessive universal  
 congenital alopecia";  
 RT Hum. Mol. Genet. 7:1671-1679(1998).  
 RL [4]  
 CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO  
 CC REGULATE ONE OF THE PHASES OF HAIR GROWTH.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKIN.  
 CC -1- DISEASE: DEFECTS IN HR ARE THE CAUSE OF ALOPECIA UNIVERSALIS  
 CC (ALONC); A RARE AUTOSOMAL RECESSIVE FORM OF HAIR LOSS. IT IS  
 CC CHARACTERIZED BY HAIR FOLLICLES WITHOUT HAIR.  
 CC -1- DISEASE: DEFECTS IN HR ARE THE CAUSE OF ATRICHA WITH PAPULAR  
 CC LESIONS (APL) (ALSO KNOWN AS CONGENITAL ATRICHA). THIS AUTOSOMAL  
 CC RECESSIVE DISEASE IS CHARACTERIZED BY PAPULAR LESIONS OVER MOST  
 CC OF THE BODY AND ALMOST COMPLETE ABSENCE OF HAIR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF039196; AAC32258.1; -  
 DR MIM; 602302; -  
 DR MIM; 203655; -  
 DR MIM; 209500; -  
 KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;  
 KW Metal-binding; Alternative splicing; Disease mutation.  
 FT ZN-FING 600 625  
 FT VARIANT 620 620 R -> Q (IN APL).  
 FT T -> A (IN ALONC).  
 FT V -> D (IN ALONC).  
 FT VARIANT 1022 1022 /FTID=VAR\_005265.  
 FT VARIANT 1136 1136 /FTID=VAR\_005266.  
 FT SEQUENCE 1189 AA; 127396 MW; 94C8B07AC8AFEE36 CRC64;  
 FT

Query Match 6.7%; Score 82.5; DB 1; Length 1189;  
 Best Local Similarity 23.3%; Pred. No. 9.5;  
 Matches 71; Conservative 23; Mismatches 92; Indels 119; Gaps 15;

OY 3 LRAGDSW--GMLACLTVMHLPAPALNRTPGPGP-----SI 40  
 DB 818 LRAGPGLRKGGLGPSVPRRLPPGALLMLDPPCPGRGHLCOEHLKKKPPVAVSGI 877  
 OY 41 OKTYDRLRYLEHQLRSIAGTY--LNYLGP-----FNEEDFNPV-----77  
 DB 878 QRLQGNLMQMTETALGALGGVQALSPGPPPSLSTFEWEGFVPELRPKSDGCVLL 937  
 OY 78 ---RLGAEPLR-----ATYDLEWMSLNDKRLT-----ON AY-----110  
 DB 938 LHRALGDEDESRVENLAASLPPEYCALHGKMLASYPGLALRPLHQLMNAIVSPH 997  
 OY 111 -SH-----LCYLRGLNUNQATATAEIRSLAHFCTSLQGLGSLAGVNAALGYPLPPLPCT 165  
 DB 998 RGLGCTKNLC-----VEADLVSLIVADPPL-----1024  
 OY 166 EPTWTPGPAHSDFLOKMD--DFWLLELOTWLR--SAKDNLKKKPPVAAVTLA 219  
 DB 1025 -PAW--HRAQKDFSLGIDGGLWSPGSQSVTWVHVRADQRIRRFJQWVTPAGAGALE 1081  
 OY 220 LGANG 224  
 DB 1082 PGAGG 1086  
 RESULT 12  
 ID P85B\_BOVIN STANDARD; PRT; 724 AA.  
 AC P23726;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 30, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE PHOSPHATIDYLINOSITOL 3-KINASE REGULATORY BETA SUBUNIT (P85-BETA).  
 DE P85-BETA SUBUNIT (PTDINS-3-KINASE P85-BETA).  
 GN PIK3R2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 CC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91191567; PubMed=1707345;  
 RA Otsu M., Hiles I.D., Gout I., Fry M.J., Ruiz-Iarrea F., Panayotou G.,  
 RA Thompson A., Dhand R., Hsuan J., Totty M., Smith A.D., Morgan S.J.,  
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;  
 RT "Characterization of two 85 kd proteins that associate with receptor  
 RT tyrosine kinases, middle-T/pp60c-src complexes, and p13-kinase";  
 RT Cell 65:91-104(1991).  
 RN [2]  
 RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.  
 RX MEDLINE=93049176; PubMed=1330535;  
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski D., Dhand R.,  
 RA Fry M.J., Blundell T.L., Wolmer A., Waterfield M.D.;  
 RT "Interaction of the p85 subunit of PI 3-kinase and its structural features  
 RT domain with a PDGF receptor phosphorylation site: structural features  
 RT and analysis of conformational changes";  
 RL EMBO J. 11:4261-4272(1992).  
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN KINASE,  
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING  
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.  
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)  
 CC SUBUNITS.  
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 RHO GAP DOMAIN.  
 CC -1- SIMILARITY: HIGH, WITH OTHER P85-BETA SUBUNIT, AND A P85 (REGULATORY)  
 CC P85-ALPHA SUBUNIT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M61746; AAA79510.1; -  
DR HSSP: P23727; 2PNA.  
DR InterPro: IPR001720; PI3K\_P85.  
DR InterPro: IPR000198; RhogAP.  
DR InterPro: IPR000980; SH2.  
DR InterPro: IPR001452; SH3.  
DR Pfam: PF006620; RhogAP\_1.  
DR Pfam: PF00017; SH2\_2.  
DR Pfam: PF00018; SH3\_1.  
DR PRINTS: PRO0401; SH2DOMAIN.  
DR PRINTS: PRO0678; PI3KINASEP85.  
DR SMART: SM00324; RhogAP\_1.  
DR SMART: SM00325; SH2\_2.  
DR SMART: SM00326; SH3\_1.  
DR PROSITE: PSS0001; SH2\_2.  
DR PROSITE: PSS0002; SH3\_1.  
DR SH3 domain; SH2 domain.  
KW DOMAIN 4 80 SH3.  
FT DOMAIN 125 255 RHO-GAP.  
FT DOMAIN 326 421 SH2\_1.  
FT DOMAIN 618 712 SH2\_2.  
SQ SEQUENCE 724 AA; 81059 MW; 902BA8B6DB087098 CRC64;

Query Match 6.7%; Score 82; DB 1; Length 724;  
Best Local Similarity 21.9%; Pred. No. 5.6;

Matches 66; Conservative 25; Mismatches 71; Indels 140; Gaps 15;

QY 6 GDSWGLACTLVMLHPVPLN-RTGDPPGSPSIQTYDITRYLEHQLSLAGTYLNY 64  
DB 42 GVAEGNERQPSVGM-MPLNERTROGDFP-----CTIYEEF 77  
QY 65 LGP-----PNEPFPNP-----R 78  
DB 78 LCPVALRGPGRPRPLPARRRCPPEGLTLPDLEQSSPDVAPILVKLEAIER 137  
QY 79 LGAETL---PRT-----VDLEWRSINDKILRTONYEAYSHLLCYLRGNROAATFEL 129  
DB 138 TGLDSYREPRVAVRTDMSISDVEOW---DAALSDGVKGF--LLALPAPLVTEPAALAA 191  
QY 130 RSL-----AHFCTSLQGL---LGSFAG-----VMALGF----- 156  
DB 192 HRALEAGPVPALPFRPTLPRLHALLTLRLQLHGLRAGRAPAPGPAVRALGATFGPPL 251  
QY 157 -----PLPQPLDGTETPTWT-PGPAHSDFLQKMDDFWLKELQDTWLMRSADKDFNRKKMKOP 211  
DB 252 LRAPPSPSPPGAGADGTEPTP---DFPALLVETKLEQHE-----EQEVAP 295  
QY 212 PA 213  
DB 296 PA 297

RESULT 13  
YM32\_MYCTU  
ID YM32\_MYCTU STANDARD; PRT; 291 AA.  
AC 010515: 010516:  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOHETICAL 30.7 KDA PROTEIN RV2232/RV2233.  
GN RV2232/RV2233 OR MT2292 OR MTCV427.13/MTCV427.14.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV

RX MEDLINE:98295987; PubMed:9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. 3rd, Tekaia F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).

RN SEQUENCE FROM N.A.  
RP STRAIN=CDC 1551 / Oshkosh;  
RC Fleschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., Deboy R., Dodson R., Gilm M.L., Holt D., H. Key E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Zitzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khoult H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases  
CC -1 SIMILARITY: BELONGS TO THE CBBY/CBBZ/GPH/YIEM FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z70692; CA94666.1; ALT\_TERM.  
DR EMBL: Z70692; CA94655.1; ALT\_INIT.  
DR TIGR: MT2292;  
DR Tuberculist; RV2232;  
DR Tuberculist; RV2233;  
KW Hypothetical protein: Complete proteome.  
SQ SEQUENCE 291 AA; 30694 MW; 750F090FB1546E5 CRC64;

Query Match 6.6%; Score 81; DB 1; Length 291;  
Best Local Similarity 26.4%; Pred. No. 2.2;  
Matches 48; Conservative 15; Mismatches 67; Indels 52; Gaps 11;

QY 23 PAVPALNRIGDPPGSPSIQTYDITRYLEHQLSLAGTY---LNYLGPVFNLYA --- 74  
DB 64 PPRAAARRRSPSPSPOL-VTFDLDTLDSARGIVSSFRNALNHGAPVPEGDLATHIV 122  
QY 75 NPPRLGAEPLP-----RATVDEWRSINDKILRTONYEAYSHLLCYLR- 118  
DB 123 GPPM--HETLRAMGGESEAEIVARADYSARGW-----AMNSLFGELPLADRT 173  
QY 119 -GLNQAAATF-----LRSLAHP-----CTSLQGLLSIAGVNA-ALGY--PLPQ 160  
DB 174 AGVRLAVATSKAEPFARILRHFGIEQHFEVYAGASTDGRSKVDVLAHLAQLRPLPE 233  
QY 161 PL 162  
DB 234 RL 235

RESULT 14  
HUPK\_RHOCA  
ID HUPK\_RHOCA STANDARD; PRT; 294 AA.  
AC P30797;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYDROGENASE EXPRESSION/FORMATION PROTEIN HUPK.  
GN HUPK.

OS Rhodobacter capsulatus (Rhodospirillum rubrum).

OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OC Rhodobacter.

ON NCBI\_TaxID=1061.

RP SEQUENCE FROM N.A.

RX MEDLINE=9177833; PubMed=2007559;

RA Xu H.W., Wall J.D.;

RT "Clustering of genes necessary for hydrogen oxidation in Rhodobacter capsulatus";

RT J. Bacteriol. 173:2401-2405(1991).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-B10;

RX MEDLINE=93268090; PubMed=8497190;

RA Colbeau A., Richard P., Toussaint B., Caballero F.J., Elster C., Delphin C., Smith R.L., Chabert J., Vignals P.M.;

RT "Organization of the genes necessary for hydrogenase expression in Rhodobacter capsulatus. Sequence analysis and identification of two hyp regulatory mutants";

RT Mol. Microbiol. 8:15-29(1993).

RL [1]

CC -1- SIMILARITY: BELONGS TO THE HUPK FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M5089; AAA2924.1; -

DR EMBL: Z15089; CAA78803.1; -

DR PIR: B38532; B38532.

DR PIR: S25691; S25691.

FT CONFLICT 70 70 A -> R (IN REF. 2).

SO SEQUENCE 294 AA; 30222 MW; 2669E8B70AF35EDA CRC64;

Query Match 6.6%; Score 80.5; DB 1; Length 294;

Best Local Similarity 30.5%; Pred. No. 2.5;

Matches 39; Conservative 3; Mismatches 39; Indels 47; Gaps 6;

QY 76 PRLGAEITLPRATVDLEWRSNDKRLTQNEYASHLYRGILNQATAE--LRRS 132

DB 36 PQQVVELLPKRL-----LFNLGCAAGGHAARLAL-----GLPAEAPAREILRDH 80

QY 133 LAHFCSTGLGSLGSGVMAALGYPLPQPLPGTEPTTPGPAHSDFLQKMDPWLKELQ 192

DB 81 LAKLCLIPKLLGL-----APQPLP-----EHWAEGGA-----ALQ 111

QY 193 TWLMSAK 200

DB 112 HMLMGAK 119

RESULT 15

BCAL\_HUMAN STANDARD; PRT; 870 AA.

ID P56945;

AC 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE CRK-ASSOCIATED SUBSTRATE (PI30CAS) (BREAST CANCER ANTI-ESTROGEN RESISTANCE 1 PROTEIN).

GN BCAR1 OR CRKAS OR CAS.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI\_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Breast cancer;

RX MEDLINE=20119429; PubMed=10639512;

RA Brinkman A., van der Flier S., Kok E.M., Dorssers L.C.;

RT "BCAR1, a human homologue of the adapter protein p130CA induces anti-estrogen resistance in breast cancer cells.";

RT J. Natl. Cancer Inst. 92:112-120(2000).

RL [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis.

RA Otto E., Birbaum S., Verbeek M., Hildebrandt F.;

RT "Interaction between human Crk-associated substrate (p130CAS) and nephrin.";

RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RL [1]

CC -1- FUNCTION: DOCKING PROTEIN WHICH PLAYS A CENTRAL COORDINATING ROLE FOR TYROSINE-KINASE-BASED SIGNALING RELATED TO CELL ADHESION. IMPLICATED IN INDUCTION OF CELL MIGRATION. OVEREXPRESSION CONFERES ANTIESTROGEN RESISTANCE ON BREAST CANCER CELLS.

CC -1- SUBUNIT: FORMS COMPLEXES IN VIVO WITH FOCAL ADHESION KINASE 1, ADAPTER PROTEIN CRKL AND LYN KINASE. CAN HETERODIMERIZE WITH CASL (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: FOCAL ADHESIONS. UNPHOSPHORYLATED FORM LOCALIZES IN THE CYTOPLASM AND CAN MOVE TO THE MEMBRANE UPON TYROSINE PHOSPHORYLATION (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH AN ABUNDANT EXPRESSION IN THE TESTIS. LOW LEVEL OF EXPRESSION SEEN IN THE LIVER, THYMUS, AND PERIPHERAL BLOOD LEUCOCYTES. THE PROTEIN HAS BEEN DETECTED IN A B-CELL LINE.

CC -1- DOMAIN: CONTAINS A CENTRAL DOMAIN (SUBSTRATE DOMAIN) CONTAINING MULTIPLE POTENTIAL SH2-BINDING SITES AND A C-TERMINAL DOMAIN CONTAINING A DIVERGENT HELIX-LOOP-HELIX (HLH) MOTIF. THE SH2-BINDING SITES, PUTATIVELY BIND CRK, NCK AND ABL SH2 DOMAINS. THE HLH MOTIF IS ABSOLUTELY REQUIRED FOR THE INDUCTION OF PSEUDOPHYAL GROWTH IN YEAST AND MEDIATES HETERODIMERIZATION WITH CASL (BY SIMILARITY).

CC -1- DOMAIN: A SERINE-RICH REGION PROMOTES ACTIVATION OF THE SERUM RESPONSE ELEMENT (SRE).

CC -1- PTM: FOCAL ADHESION KINASE 1 PHOSPHORYLATES THE PROTEIN AT THE YDYLH MOTIF. THE SRC FAMILY KINASES ARE RECRUITED. THE PHOSPHORYLATED SITES AND CAN PHOSPHORYLATE OTHER TYROSINE RESIDUES. TYROSINE PHOSPHORYLATION IS TRIGGERED BY INTEGRIN-MEDIATED ADHESION OF CELLS TO THE EXTRACELLULAR MATRIX.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN. NECESSARY FOR THE LOCALISATION OF THE PROTEIN TO FOCAL ADHESIONS AND INTERACTS WITH ONE PROLINE-RICH REGION OF FOCAL ADHESION KINASE 1.

CC -1- SIMILARITY: CONTAINS 1 SH3-BINDING DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE CAS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AJ242987; CAB75875.2; -

DR EMBL: AF218451; AAF27527.1; -

DR MIM: 602941; -

DR InterPro: IPR001452; SH3.

DR Pfam: PF00018; SH3.1.

DR SMART: SM00326; SH3.1.

DR PROSITE: PS50002; SH3.1.

KW Phosphorylation; SH3 domain;

FT DOMAIN 3 65.

FT DOMAIN 74 87.

FT DOMAIN 115 416.

FT DOMAIN 422 614.

FT DOMAIN 635 643.

FT DOMAIN 746 796.

FT DOMAIN 766 796.

FT CONFLICT 236 236.

FT CONFLICT 349 349.

FT CONFLICT 363 363.

SO SEQUENCE 870 AA; 93361 MW; D57A9CACA4FAD0 CRC64;



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:30:24 ; Search time 75.65 seconds  
(without alignments)  
435.047 Million cell updates/sec

Title: US-09-931-704-2  
Perfect score: 1226  
Sequence: 1 MDLRAGDSWGMACICTYLVW.....KKMQPPAAVTLHGAHGF 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_17: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mmc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_protist: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1226	100.0	225	4 Q9UBD9	Q9ubd9 homo sapien
2	1193	97.3	225	4 Q9QZM3	Q9qzm3 mus musculu
3	1307.5	12.3	215	13 Q9PUJ2	Q9puj2 plethodon j
4	1507.5	12.3	215	13 Q9PUJ1	Q9puj1 plethodon j
5	1507.5	12.3	215	13 Q9PUJ0	Q9puj0 plethodon j
6	148.5	12.1	215	13 Q9PUJ9	Q9puj9 plethodon j
7	96.5	7.9	455	11 Q9CWW7	Q9cww7 mus musculu
8	93	7.6	559	4 Q99799	Q99799 homo sapien
9	93	7.6	619	4 Q9UOL2	Q9uol2 homo sapien
10	93	7.6	619	4 Q9UBU9	Q9ubu9 homo sapien
11	92	7.5	332	10 Q9MAU1	Q9mau1 arabidopsis
12	92	7.5	423	11 Q9JHE4	Q9jhe4 mus musculu
13	90	7.3	423	11 Q9DBV6	Q9dbv6 mus musculu
14	89.5	7.3	771	2 Q9S3Q9	Q9s3q9 porphyromon
15	89	7.2	733	2 Q91664	Q91664 pseudomonas
16	88.5	7.2	1429	4 Q9HAV8	Q9hav8 homo sapien
17	87	7.1	955	11 Q88287	Q88287 mus musculu
18	87	7.1	1561	11 Q88286	Q88286 mus musculu
19	86.5	7.1	640	2 Q54153	Q54153 streptomyce

20	85.5	7.0	1154	11 Q921R2	Q921r2 mus musculu
21	85	6.9	727	11 Q88841	Q88841 mus musculu
22	84.5	6.9	294	11 Q9CPZ1	Q9cpz1 mus musculu
23	84.5	6.9	389	11 Q91A71	Q91a71 aeryopyrum p
24	84	6.9	294	4 Q9NXX5	Q9nxx5 homo sapien
25	84	6.9	411	2 Q9RVP3	Q9rvp3 delinococcus
26	84	6.9	559	2 Q83432	Q83432 treponema p
27	84	6.9	876	4 Q9P210	Q9p210 homo sapien
28	83.5	6.8	310	4 Q9S159	Q9s159 homo sapien
29	83.5	6.8	5990	2 Q9RLP6	Q9rlp6 mycobacteri
30	83	6.8	222	2 Q9LS84	Q9ls84 streptococc
31	83	6.8	625	4 Q16205	Q16205 homo sapien
32	83	6.8	1620	4 Q9Y4K6	Q9y4k6 homo sapien
33	82.5	6.7	310	4 Q9UID0	Q9uid0 homo sapien
34	82.5	6.7	542	2 Q9HTD5	Q9htd5 pseudomonas
35	82.5	6.7	1189	4 Q9NPE1	Q9npe1 homo sapien
36	82	6.7	283	2 Q92B87	Q92b87 pseudomonas
37	82	6.7	1179	10 Q65507	Q65507 arabidopsis
38	82	6.7	2376	10 Q9F1N7	Q9f1n7 arabidopsis
39	81.5	6.6	396	2 Q9AF00	Q9af00 frankia sp.
40	81	6.6	328	12 Q9E125	Q9e125 human immun
41	81	6.6	328	12 Q9E122	Q9e122 human immun
42	81	6.6	618	11 Q88984	Q88984 rattus norv
43	80.5	6.6	924	2 Q9X582	Q9x582 rhodothermu
44	80.5	6.6	305	4 Q9H7R6	Q9h7r6 homo sapien
45	80.5	6.6	385	2 Q9AD13	Q9ad13 streptomyce

## ALIGNMENTS

RESULT	ID	Q9UBD9	PRELIMINARY:	PRT;	225 AA.
AC	Q9UBD9	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)				
DE	NEUTROPHILIN-1/B-CELL STIMULATING FACTOR-3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homn.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99432254; PubMed=10500198;				
RA	Senaldi G., Varum B.C., Sarmiento U., Starnes C., Lile J., Scully S.,				
RA	Guo J., Elliott G., McNinch J., Shaklee C.L., Freeman D., Mann F.,				
RA	Simonet W.S., Boone T., Chang M.-S.;				
RT	"Novel neurotrophin-1/B cell-stimulating factor-3: A cy... line of the				
RT	IL-6 family.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 96:11458-11463(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99382254; PubMed=10448081;				
RA	Shi Y., Wang W., Yourey P.A., Gohari S., Zukauskas D., Zhang J.,				
RA	Ruben S., Alderson R.F.;				
RT	"Computational EST database analysis identifies a novel member of the				
RT	neuropoietic cytokine family.";				
RL	Biochem. Biophys. Res. Commun. 262:132-138(1999).				
DR	EMBL: AF176912; AAF00992.1; -				
DR	EMBL: AF172854; AAD54284.1; -				
DR	EMBL: AF176911; AAF00991.1; -				
SQ	SEQUENCE 225 AA; 25176 MW; E2DDAB6280833B55 CRC64;				

Query Match 100.0%; Score 1226; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.5e-109; Indel 0; Gaps 0;  
Matches 225; Conservative 0; Mismatches 0;

QY 1 MDLRAGDSWGMACICTYLVWHPAIPALNRGDPGPGSIQKTYDLYRLQLRSLAGT 60  
|||||  
DB 1 MDLRAGDSWGMACICTYLVWHPAIPALNRGDPGPGSIQKTYDLYRLQLRSLAGT 60

QY 61 YLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLCYRLG 120  
 DB 61 YLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLCYRLG 120  
 QY 121 NROAAATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWTGPAHSDFIQ 180  
 DB 121 NROAAATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWTGPAHSDFIQ 180  
 QY 181 KMDDFWLKELQTLWMSAKDFNRLKKKQPPAAATVTLHGANGF 225  
 DB 181 KMDDFWLKELQTLWMSAKDFNRLKKKQPPAAATVTLHGANGF 225

RESULT 2  
 ID Q9QZM3 PRELIMINARY: PRT: 225 AA.  
 AC Q9QZM3;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE NEUROKOTOPHIN-1/B-CELL STIMULATING FACTOR-3.  
 GN BSF3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99432254; PubMed=10500198;  
 RA Senaldi G., Varnum B.C., Sarmiento U., Starnes C., Lile J., Scully S.,  
 RA Guo J., Elliott G., McIninch J., Shaklee C.L., Freeman D., Mann F.,  
 RA Simonet W.S., Boone T., Chang M.-S.;  
 RT "Novel neurotrophin-1/B cell-stimulating factor-3: A cytokine of the  
 RT IL-6 family."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11458-11463(1999).  
 DR EMBL: AF176913; AAF00993.1;  
 DR MGD: MGI:1930088; Bsf3.  
 SQ SEQUENCE 225 AA; 25261 MW; 68B1FEAAB7F1A950 CRC64;

Query Match 97.3%; Score 1193; DB 11; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 1.2e-105;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MDLRAGDSMGWLACTCYLWHLPAVPAALNRPDGPSPSTOKTYDRLRYLEHQRSLAGT 60  
 DB 1 MDLRAGDSMGWLACTCYLWHLPAVPAALNRPDGPSPSTOKTYDRLRYLEHQRSLAGT 60  
 QY 61 YLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLCYRLG 120  
 DB 61 YLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLCYRLG 120  
 QY 121 NROAAATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWTGPAHSDFIQ 180  
 DB 121 NROAAATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWTGPAHSDFIQ 180  
 QY 181 KMDDFWLKELQTLWMSAKDFNRLKKKQPPAAATVTLHGANGF 225  
 DB 181 KMDDFWLKELQTLWMSAKDFNRLKKKQPPAAATVTLHGANGF 225

RESULT 3  
 ID Q9PUJ2 PRELIMINARY: PRT: 215 AA.  
 AC Q9PUJ2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE RECEPTIVITY FACTOR ISOFORM 1 PRECURSOR.  
 GN PRF.  
 OS Plethodon jordanii (Salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Caudata; Salamandroidea; Plethodonti  
 OC Plethodon.  
 OX NCBI\_TaxID=8336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99420364; PubMed=10489368;  
 RA Rollmann S.M., Houck L.D., Feldhoff R.C.;  
 RT "Proteinaceous pheromone affecting female receptivity in  
 RT salamander."  
 RL Science 285:1907-1909(1999).  
 DR EMBL: AF181480; AAF01025.1;  
 FT Signal.  
 SQ SEQUENCE 215 AA; 24138 MW; B1906BB66335738 CRC64;

Query Match 12.3%; Score 150.5; DB 13; Length 215;  
 Best Local Similarity 26.2%; Pred. No. 1.5e-06;  
 Matches 42; Conservative 31; Mismatches 80; Indels 7; Gaps 3;  
 QY 56 SLAGTYLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLC 115  
 DB 56 SLAGTYLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLC 115  
 QY 116 YLR-GLNRQ----AATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWT 170  
 DB 116 YLR-GLNRQ----AATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWT 170  
 QY 171 PGPAHSDFIQKMDDFWLKELQTLWMSAKDFNRLKKKQ 210  
 DB 171 PGPAHSDFIQKMDDFWLKELQTLWMSAKDFNRLKKKQ 210  
 QY 173 PEGSAVFRKKLGGVYCKEYKERVLLTKRDFELAKKQ 212  
 DB 173 PEGSAVFRKKLGGVYCKEYKERVLLTKRDFELAKKQ 212

RESULT 4  
 ID Q9PUJ1 PRELIMINARY: PRT: 215 AA.  
 AC Q9PUJ1;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE RECEPTIVITY FACTOR ISOFORM 2 PRECURSOR.  
 GN PRF.  
 OS Plethodon jordanii (Salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Plethodonti; Pleth.  
 OX NCBI\_TaxID=8336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99420364; PubMed=10489368;  
 RA Rollmann S.M., Houck L.D., Feldhoff R.C.;  
 RT "Proteinaceous pheromone affecting female receptivity in a terrestrial  
 RT salamander."  
 RL Science 285:1907-1909(1999).  
 DR EMBL: AF181481; AAF01026.1;  
 FT Signal.  
 SQ SEQUENCE 215 AA; 24080 MW; B341BBB7B4E28438 CRC64;

Query Match 12.3%; Score 150.5; DB 13; Length 215;  
 Best Local Similarity 26.2%; Pred. No. 1.5e-06;  
 Matches 42; Conservative 31; Mismatches 80; Indels 7; Gaps 3;  
 QY 56 SLAGTYLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLC 115  
 DB 56 SLAGTYLNTLGPPFNPDPNPRLGAETLPRAVDLEWRSINDKRLTQNEAYSHLLC 115  
 QY 116 YLR-GLNRQ----AATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWT 170  
 DB 116 YLR-GLNRQ----AATLRLSLAHFCTSLQGLGSIAGVMAALGYPLPPLPTEPTWT 170  
 QY 171 PGPAHSDFIQKMDDFWLKELQTLWMSAKDFNRLKKKQ 210  
 DB 171 PGPAHSDFIQKMDDFWLKELQTLWMSAKDFNRLKKKQ 210



D6	173	PFECSAYFRKKLRGCVCKEYKERVLLTKRDFEEFLAKKYQ	212
<hr/>			
RESULT	5		
O9PUJ0			
ID	O9PUJ0	PRELIMINARY;	PRT; 215 AA.
AC	O9PUJ0:		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)		
DE	RECEPTIVITY FACTOR ISOFORM 3 PRECURSOR.		
GN	PRF.		
OS	Plethodon jordani (Salamander).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;		
OC	Plethodon.		
OX	NCBI_TaxID=8336;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99420364; PubMed=10489368;		
RA	Rollmann S.M., Houck L.D., Feldhoff R.C.;		
RT	"Proteinaceous pheromone affecting female receptivity in a terrestrial		
RL	salamander." Science 285:1907-1909(1999).		
DR	EMBL; AF181482; AAF01027.1; -.		
KW	Signal.		
FT	SIGNAL.		
SO	SEQUENCE	1 23 POTENTIAL. 215 AA; 24024 MW; 6EB6E64AD93A1343 CRC64;	
<hr/>			
Query Match	12.3%;	Score 150.5;	DB 13; Length 215;
Best Local Similarity	27.1%;	Pred. NO. 1.5e-06;	
Matches	45; Conservative	34; Mismatches	72; Indels 15; Gaps
OY	56 SLAGTYLANLGPPFNPDNPRLGAETLPRAATVDLEWRSINDKRLTQNTEAVSHLIC	115	
DB	55 SLPLPTLSOGAPLSPDQDLPHIKVNLPTAMODVTRQTDERRLNKNLYFSAIYE	114	
OY	116 YLR-----GLN--RQAATAEIIRSLAHFCISLOGLCISAGWAALCYPLPPIPGTE	166	
DB	115 FLEAMTEOEDINPAELALKAKEFEEMANSNT---LISKISDIMQMGSVTITLP--K	168	
OY	167 PWMTPEPAHSDFLDKMDDEFMLKELOTWTLMRSKDKDNFLKKMKOP	212	
DB	169 PLVVPKSGSAFSSKLRGGVCKEYKERVLTTRDMMLAEKYQP	214	
<hr/>			
RESULT	6		
O9PUJ0			
ID	O9PUJ0	PRELIMINARY;	PRT; 215 AA.
AC	O9PUJ0:		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)		
DE	RECEPTIVITY FACTOR ISOFORM 4 PRECURSOR.		
GN	PRF.		
OS	Plethodon jordani (Salamander).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;		
OC	Plethodon.		
OX	NCBI_TaxID=8336;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99420364; PubMed=10489368;		
RA	Rollmann S.M., Houck L.D., Feldhoff R.C.;		
RT	"Proteinaceous pheromone affecting female receptivity in a terrestrial		
RL	salamander." Science 285:1907-1909(1999).		
DR	EMBL; AF181483; AAF01028.1; -.		
KW	Signal.		
FT	SIGNAL.		
SO	SEQUENCE	1 23 POTENTIAL. 215 AA; 24054 MW; A4A412135FFA4E7F CRC64;	

	Query Match	12.1%	Score 148.5;	DB 13:	Length 215;	
	Best Local Similarity	25.9%	Pred. No. 2,3e-06;			
	Matches 42;	Conservative 33;	Mismatches 80;	Indels 7;	Gaps 3;	
Oy	SLAGTYLNTYLPPEPPEDNPRLGAETLLPRATVDELYWRSINDKRLITYONYEAYSILLC	115				
Dd	55 SLPLTFSSFGCAISPDPYRPHIKYVNLPPAAADYDFPIQOTETRLKNM;PYSAIVE	114				
Oy	116 YLR-GLANQ-----ATAALRRSLAHFCISLOGLGSLNGVAAGLYPLPQL;TWTEPTWT	170				
Dd	115 FLKEAMTQEDLINPAEISLKAKFEERAMANSNTLSISKIDIMOMGMSVTITIF--KLPLV	172				
Oy	171 PGPAHSDELQKMDPFMLKELOTWLMRSACFNRLKKRMOP	212				
Dd	173 PPKGSAYFSKRLRGVGVCKEYKEREFLLRKDRDMLLAERYOGP	214				
RESULT						
ID	09CWN7	PRELIMINARY;	PRT,	455 AA.		
AC	09CWN7					
DT	01-JUN-2001 (TREMBLrel. 17, Created)					
DT	01-JUN-2001 (TREMBLrel. 17, last sequence update)					
DT	01-JUN-2001 (TREMBLrel. 17, last annotation update)					
DE	2410003H12RIK. PROTEIN.					
GN	2410003H12RIK.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.					
RN	NCBI_Taxid=10090;					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=C57BL/6J; TISSUE=EMBRYONIC STEM CELLS;					
RX	MEDLINE=21085660; PubMed=11217851;					
RA	Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,					
RA	Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,					
RA	Aizawa K., Izawa M., Nishi K., Kiyooswa H., Kondo S., Yamanka I.,					
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,					
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,					
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,					
RA	Kuehl P., Lewis S., Matsuo Y., Miyai D.I., Pesole G., Quackenbush J.,					
RA	Schmili L.M., Staubli F., Suzuki R., Tomita M., Wagner I., Washio T.,					
RA	Sakai K., Offelli D., Furuno M., Anou H., Baldarelli R., Barsh G.,					
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Galiboldi M.,					
RA	Gustincich S., Hill D., Hofman M., Humé D.A., Kamiya M., Lee N.H.,					
RA	Lyon P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,					
RA	Norodine P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,					
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,					
RA	Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitteker C., Wilming L.,					
RA	Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,					
RA	Hayashizaki Y.,					
RT	"Functional annotation of a full-length mouse cDNA collection."					
RL	Nature 409:685-690(2001).					
DR	EMBL; AK010358; BAB26878.1; ;					
MD	MG1:1919221; 2410003H12RIK.					
SQ	SEQUENCE 455 AA; 51481 MW; 201886B814EB9CFE CRC64;					
Query Match	7.9%; Score 96.5; DB 11: Length 455;					
Best local similarity	22.4%; Pred. No. 0.5;					
Matches 59;	Conservative 24;	Mismatches 89;	Indels 91;	Gaps 12		
Oy	7 DSMGMCLACTIVM-HLPAVPALNRGGDG-----PGSIOCTYDITTRYLEHQLS	56				
Dd	29 DANYNOSCFFMODLFAPHPVEVLFTIRHTYRNPGKVLYFLVDGRPNVOYCECPARAV-----	82				
Oy	57 LAGTYLVNYIGAPPENE-----PDENPRLGAETLLPRATVDELYWRSINDKRLTL	104				
Dd	83 -----YFAIPTNEDARGIAQMFOVFCKRNP-----AWERNYNTIL--- 116					

QY 105 QNENAVSHLLCYLGLNROQATAEIARSLAHFCTSLGGLGSIAGVMAALGYPPLPOP--- 161  
 DB 117 -----VDPHFL-LLEPTLMEPTAEVLLSAFHICKFLOKRYQL-----PLEGQVQR 162  
 QY 162 -----LPGTEPTWTPRGPAH-----SDFLOKMDDFLLKELOTWL---WRSKADF 202  
 DB 163 LLSSLSLOSTWCSATAGMLRKLYTLNNCIDSSRLPELHSHWLLND-RTWLAHRRSRAOS 221  
 QY 203 NRLKKMQPPAAAVTLHGAHGF 225  
 DB 222 SRYFQSLIMAHILISQFFGTTF 244

RESULT 8  
 Q99799 PRELIMINARY; PRT; 559 AA.  
 ID Q99799  
 AC Q99799;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DE TIP ASSOCIATING PROTEIN.  
 GN TAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97318898; PubMed=9175835;  
 RA Yoon D.W., Lee H., Seol W., Demaria M., Rosenzweig M., Jung J.U.;  
 RT "Tap: a novel cellular protein that interacts with tip of herpesvirus  
 satelit and induces lymphocyte aggregation.";  
 RL EMBL: U60073; AAB8111.1; -  
 DR InterPro: IPR001611; LRR.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00446; LRRcap. 1.  
 SQ SEQUENCE 559 AA; 63301 MW; 250C136C62042C0C CRC64;

Query Match 7.6%; Score 93; DB 4; Length 559;  
 Best Local Similarity 24.7%; Pred. No. 1.4;  
 Matches 60; Conservative 32; Mismatches 93; Indels 58; Gaps 13;

QY 5 AGDSWML-----ACLCYVHLPAVPAALNRTGDP-GGEPSTOKTYDIT--RYLEHQL 54  
 DB 340 SDDRQGLDAYHDGACCSLSIPFIQNPARRSLAEYFKDSRNKLLKDPFLRLKHTR 399  
 QY 55 RSLAGTYLNYLGPFPNPPDPNP--PRLGAETLPRAIVDL-----EYWRSLNDKLR-LTON 106  
 DB 400 LNVV-AFLNEL--PKQHVNVSFYVDISQSTSLCFSVNGVKEVDGKSRSLSRAFTTR 456  
 QY 107 Y---EAYSHLLCYLRG--LNROQATAEIARSLAHFCTSLGGLGSIAGVMAALGYPPLPOP 161  
 DB 457 FIAVPASNSGICIVNDELFRNASSEIQRIFA-----MPAP 493  
 QY 162 LPGEPTWTPRGPAHSDFLQK-----MDDFWLKELOTWL---RSKDPNRLKKMQPP 212  
 DB 494 TPSSSPVPTLSPQOEMLQAFSTQSGMNLMSQKCLQDNNMDYTRSAQAFTHLKAKEIP 553  
 QY 213 AAA 215  
 DB 554 EVA 556

RESULT 9  
 Q990L2 PRELIMINARY; PRT; 619 AA.  
 ID Q990L2  
 AC Q990L2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE TIP-ASSOCIATED PROTEIN TAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zolotukhin A.S., Tan W., Bear J., Tabernero C., Felder B.K.;  
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF126246; AAD20016.1; -  
 DR InterPro: IPR001611; LRR.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00446; LRRcap. 1.  
 SQ SEQUENCE 619 AA; 70195 MW; 092D38037610BF3A CRC64;

Query Match 7.6%; Score 93; DB 4; Length 619;  
 Best Local Similarity 24.7%; Pred. No. 1.6;  
 Matches 60; Conservative 32; Mismatches 93; Indels 58; Gaps 13;

QY 5 AGDSWML-----ACLCYVHLPAVPAALNRTGDP-GGEPSTOKTYDIT--RYLEHQL 54  
 DB 400 SDDRQGLDAYHDGACCSLSIPFIQNPARRSLAEYFKDSRNKLLKDPFLRLKHTR 459  
 QY 55 RSLAGTYLNYLGPFPNPPDPNP--PRLGAETLPRAIVDL-----EYWRSLNDKLR-LTON 106  
 DB 460 LNVV-AFLNEL--PKQHVNVSFYVDISQSTSLCFSVNGVKEVDGKSRSLSRAFTTR 516  
 QY 107 Y---EAYSHLLCYLRG--LNROQATAEIARSLAHFCTSLGGLGSIAGVMAALGYPPLPOP 161  
 DB 517 FIAVPASNSGICIVNDELFRNASSEIQRIFA-----MPAP 553  
 QY 162 LPGEPTWTPRGPAHSDFLQK-----MDDFWLKELOTWL---RSKDPNRLKKMQPP 212  
 DB 554 TPSSSPVPTLSPQOEMLQAFSTQSGMNLMSQKCLQDNNMDYTRSAQAFTHLKAKEIP 613  
 QY 213 AAA 215  
 DB 614 EVA 616

RESULT 10  
 Q990B9 PRELIMINARY; PRT; 619 AA.  
 ID Q990B9  
 AC Q990B9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE TIP ASSOCIATING PROTEIN (NUCLEAR RNA EXPORT FACTOR 1).  
 GN TAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99219874; PubMed=10202159;  
 RA Braun I.C., Rohrbach E., Izaurralde E.;  
 RT "Tap binds to the retroviral constitutive Transport Element" through a  
 RT novel RNA-binding motif which also mediates CTE-dependent RNA export  
 RT from the nucleus.";  
 RL EMBL: J. 18:1966-1973(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kang Y., Bogerd H., Izaurralde E., Cullen B.R.;  
 RT "A novel RNA binding domain in human Tap protein is necessary but not  
 RT sufficient for the nuclear export of mRNA containing Mason-Pfizer  
 RT monkey virus constitutive transport element.";  
 RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CHORIOCARCINOMA;

RA Strausberg R.  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ132712; CAAL0753.1;  
 DR EMBL: AF112880; AAD39102.1;  
 DR EMBL: BC004504; AAH04504.1;  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003603; LRRcap.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00446; LRRcap.1.  
 SQ SEQUENCE 619 AA; 70182 MW; 338872AADA789FBF CRC64;

Query Match 7.6%; Score 93; DB 4; Length 619;  
 Best Local Similarity 24.7%; Pred. No. 1.6; Mismatches 93; Indels 58; Gaps 13;  
 Matches 60; Conservative 32; Mismatches 93; Indels 58; Gaps 13;

QY 5 AGDSMGML-----ACLTCTVLMHLPAVPALNRTGDP--GPGPSIOKTYLDT---RYLEHQL 54  
 DB 400 SGRDGLDAVHDGACCSLSIFIPONPARSSIAEYFKDSNNVKLKDPTLRFRLKHTR 459  
 QY 55 RSLAGTYLNTGPPNEDEPNP--PRIGAEPLPRATVDL-----EYWSLNDKLR-LTON 106  
 DB 460 LNVV-AFLNEL--PKTDVNSFVVDISAQSTLLCFSGVNGVFKEDSKSRDLSRAFTRT 516  
 QY 107 Y---EAVSHLCYLRG--LNROAATAELRSLAHFCTSLQGLGSLAGVMAALGYPLDOP 161  
 DB 517 FIAPVANSGLCIYNDLFLVKNASSEIQRAPA-----MAP 553  
 QY 162 LPGETPTWTPGPAHSDFLOK-----MDFFMLLKELOTWLV--RSKADFNLRKKKQPP 212  
 DB 554 TPSSPVPVTLSPDEQEMLQAFSTGSGMNLWMSQKCLQDNNMDYTRSAQAFHLKAKEIP 613  
 QY 213 AAA 215  
 DB 614 EVA 616

RESULT 11  
 ID 09MAU1 PRELIMINARY; PRT; 332 AA.  
 AC 09MAU1;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE F13M7.8 PROTEIN.  
 GN F13M7.8.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Kwan A., Liu, S.,  
 RA Li J., Kirmenetskaia I., Luros J., Araujo R., Au M., Bredel V.,  
 RA Buehler E., Conway A., Dewar K., Feng J., Kim C., Kurtz D., Li Y.,  
 RA Palm C., Shim P., Sun H., Davis R., Ecker J., Federspiel N.,  
 RA Theologis A.;  
 RT "The sequence of BAC F13M7 from Arabidopsis thaliana chromosome 1,"  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RP [4]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;  
 RA Theologis;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC004809; AAF40444.1;  
 DR InterPro: IPR002965; P-rich\_extensn.  
 DR PRINTS: PR01217; PRICHEXTENS.  
 SQ SEQUENCE 332 AA; 36793 MW; 18E8687141A070F4 CRC64;

Query Match 7.5%; Score 92; DB 10; Length 377;  
 Best Local Similarity 26.4%; Pred. No. 0.92; Mismatches 72; Indels 66; Gaps 13;  
 Matches 58; Conservative 24; Mismatches 72; Indels 66; Gaps 13;

QY 23 PAVPALNRTGDPGPPPSIOKTYDLTRYLEHQLSLAGTYLNT----- 64  
 DB 113 PSYTAGNLISGY-P-PPSP--TYDPGPEYRQWESLQOEFRRNPQIRPLPLIGLSPYG 169  
 QY 65 LGPPNEDEPNPRLIGAEPLPRATVDLEWWSLNDKLRITQWVAVSHLLCYLRGLNROA 124  
 DB 170 LGPIRASPQFLQPRVAP--PISILD-----TSRNNKAR-----SKDGLAVVRG--KKV 215  
 QY 125 ATAEIRSL-----AHFCTSLQGLGSLAGVMAALGYPLDOPLP -GTEPTWT 170  
 DB 216 RTEGSSSLYSIGRSMKNGAHV-----GIQPORSGIMK-----PLKPLPVLNLTETSV 266  
 QY 171 PGPASDFLOKMDDFMLLKELOTWLMRSKADFNLRKKKQ 210  
 DB 267 DDPDEBSADEKEDBEAVAKQL-----SEKDL--LKRRIE 298

RESULT 12  
 ID 09JHE4 PRELIMINARY; PRT; 423 AA.  
 AC 09JHE4;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE CEREBROSIDE SULFOTRANSFERASE.  
 GN CST.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20193614; PubMed=10727929;  
 RA Hirahara Y., Tsuda M., Wada Y., Honke K.;  
 RT "CDNA cloning, genomic cloning, and tissue-specific regulation of  
 RT mouse cerebroside sulfotransferase,"  
 RL Eur. J. Biochem. 267:1909-1917(2000).  
 DR EMBL: AB032940; BAA93009.1;  
 DR EMBL: AB032939; BAA93008.1;  
 DR MGD: MGI:1858277; Cst.  
 KW Transferrase.  
 KW SEQUENCE 423 AA; 48984 MW; FD4A2BAID983E46 CRC64;

Query Match 7.5%; Score 92; DB 11; Length 423;  
 Best Local Similarity 21.8%; Pred. No. 1.2; Mismatches 60; Indels 106; Gaps 12;  
 Matches 53; Conservative 24; Mismatches 60; Indels 106; Gaps 12;

QY 44 YDLTRYLEHQLSLAGTYLNT---LGGP-----INP----- 72  
 DB 204 YDPSSYNAHYLRNLFFPDGYDSSLPASPRVQEHILEVRRPHVLQGYFDESIVLLR 263  
 QY 73 -----DENPPRLIGAEPLPRAT---VDLEWWSLNDKLRITQWVAVSHLLCYLRGLNROA 109  
 DB 264 ELICWDLEDVLYFKLNARDSPVPRISGELYRATAMNLLDVLRYHFN--FEMKRVFA 321  
 QY 110 YSHLLCYLRGLNROA-ATAELRSL--LAHFCTSLQGLGSLA-----GVMAA 153  
 DB 322 F-----GREMARVAVELRQANEHMRHICIDGGQAVGAELIRDSANLIL--PLGIGKSI 373



DR EMBL: AE004483; AAC03843.1; -;  
KM Hypothetical protein; Complete proteome.  
SO SEQUENCE 733 AA; 81538 MW; 580F9DDBCB3909DB CRC64;

Query Match	7.3%	Score 89	DB 2	Length 733
Best Local Similarity	26.6%	Pred. No. 4.6		
Matches	58	Conservative	26	Mismatches 66
				Indels 68
				Gaps 13

QY	28	LNRIGDPGPSPISQKTYDLTRYLF-----HQRLSLAGTYLNLGPPFNEDP--FNPP	77
		:     :	
Db	230	LNRLGHGRGPKV-----SRILKLYFTADVDHEASSSHYPNRLAEAFPSVDLFRQ	283
		:     :	
QY	78	RL-----GAETLRPATVDLEWRSINDKLRLTQVAYEASHLCTLRGLNRAAFAELRR	131
		:    :	
Db	284	RLINQSGKACQALARA-----IRLRQFFD-YA-----DRLALUEDLQA	320
		:    :	
QY	132	SLAHF-----CTSLQGLGSIAGVMAALGYPLPQLPGTEPTWPGPAHSDFLQKMDPWL	187
		:     :    :    :	
Db	321	SLFHLRQOSNPAMWGLRLSL--GALAAANTLTLDRLKLAGA-----SNPDATADQDSAL	371
		:     :    :    :	
QY	188	LKELQTWLMRSKAD--FNRLKKMKMPPAAAVLNLGHAGH	224
		:     :     :    :	
Db	372	LDRSP-----RSLKDAFERLRQOLTP-----TSLERHG	400
		:     :     :    :	

Search completed: March 18, 2002, 09:39:15  
Job time: 531 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:28:54 : Search time 55.47 Seconds  
(without alignments)  
300.460 Million cell updates/sec

Title: US-09-931-704-2  
Perfect score: 1226  
Sequence: 1 MDLRAGDSWGLACTVLM.....KKKMPAAAVTLHGAGHF 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_1101.\*

1: /SID2/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1226	100.0	225	19	AAW29715 Human neurotrophic
2	1226	100.0	225	19	AAW56141 Amino acid sequenc
3	1226	100.0	225	20	AAW94466 Human cardiotoxiph
4	1226	100.0	225	21	AAW7813 Human NMT-1 protei
5	1226	100.0	225	22	AAW63543 Amino acid sequenc
6	1226	100.0	225	22	AAW25831 Human protein sequ
7	1214	99.0	223	22	AAW29716 Human cardiotoxiph
8	1193	97.3	225	19	AAW29716 Mouse neurotrophic
9	1193	97.3	225	19	AAW56142 Amino acid sequenc
10	1193	97.3	225	21	AAW87814 Murine NMT-1 prote
11	1169	95.4	215	21	AAW19586 Human interleukin-

12	1136	92.7	215	21	AAW19587
13	885	72.2	164	22	AAW20115
14	885	72.2	164	22	AAW54012
15	162.5	13.3	208	20	AAW09197
16	160	13.1	203	16	AAW09196
17	118.5	9.7	203	16	AAW83965
18	118.5	9.7	203	17	AAW88204
19	118.5	9.7	203	18	AAW29237
20	96.5	7.9	243	22	AAW20277
21	92	7.5	332	21	AAW22132
22	92	7.5	332	21	AAW40321
23	91.5	7.5	201	16	AAW83967
24	91.5	7.5	201	16	AAW29238
25	91.5	7.5	201	20	AAW06490
26	91.5	7.5	201	21	AAW27662
27	91.5	7.5	201	21	AAW13004
28	91.5	7.5	201	21	AAW93697
29	91.5	7.5	201	21	AAW87818
30	91.5	7.5	201	22	AAW50994
31	91	7.4	195	14	AAW34432
32	91	7.4	195	20	AAW83337
33	91	7.4	242	22	AAW20275
34	89.5	7.3	771	20	AAW34574
35	89.5	7.3	766	20	AAW34431
36	85.5	7.0	429	12	AAW10976
37	84.5	6.9	1473	22	AAW06756
38	84	6.9	200	19	AAW37931
39	83.5	6.8	310	20	AAW87993
40	83.5	6.8	399	20	AAW34543
41	83.5	6.8	450	20	AAW34409
42	82.5	6.7	582	14	AAW41000
43	82.5	6.7	1189	20	AAW15217
44	82	6.7	200	15	AAW33425
45	82	6.7	200	20	AAW83340

## ALIGNMENTS

## RESULT 1

ID AAW29715 standard; Protein; 225 AA.

AC AAW29715;

DT 09-NOV-1998 (first entry)

DE Human neurotrophic factor NMT-1.

XX NMT-1; neurotrophic factor; human; antiinflammatory; adjuvant;

KW Alzheimer's disease; Parkinson's disease; Huntington's d disease;

KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndr; neur;

KW peripheral neuropathy; dystrophy; neural retina degene; CVID; selective Iga d; deficiency;

KW common variable immunodeficiency; CVID; selective Iga d; deficiency;

KW hypogammaglobulinemia; X-linked agammaglobulinemia; au; septic;

KW therapy.

OS Homo sapiens.

XX Key

FT Peptide

FT Protein

FT W09833922-A1.

XX 06-AUG-1998.

XX 02-FEB-1998; 98W0-US02363.

XX 30-JAN-1998; 98US-0016534.

PR 03-FEB-1997; 97US-0792019.

M use interleukin-  
Protein #6549 enco  
Protein #8049 enco  
Human DNA interie  
Human DNAX interie  
Human cardiac hype  
Human cardiotoxiph  
Human cardiotoxiph  
Human cardiotoxiph  
Human interleukin  
Human interleukin  
Human thallia  
Human cardiac hype  
Human cardiac hype  
Human tumour-asso  
Human protein PRO8  
Human cardiotoxiph  
Human cardiotoxiph  
Human cardiotoxiph  
Human PRO882 prote  
Sequence of growth  
Chicken ciliary ne  
Human Interleukin  
Porphyromonas ging  
Porphyromonas ging  
Porphyromonas ging  
Human G-protein co  
Amino acid sequenc  
Human MCG4 prote  
Porphyromonas ging  
Porphyromonas ging  
Human brain CDNA c  
Human Hairless wil  
Human/rat chimeric  
Modified ciliary n







RESULT	5
AAG63543	
ID	AAG63543 standard; Protein; 225 AA
XX	
AC	

Amino acid sequence of a human NNT-1 protein.

muscular mass; paralysis; cancer; obesity; fertility; endometriosis;  
blastocyst implantation; thrombosis; retinal disease;  
retinal pigmentosus.

02-AUG-2001.

27-JAN-2000; 2000FR-0001035.  
12-OCT-2000; 2000FR-0013089.

WPI; 2001-488773/53.  
N-PSDB; AAH74484.

Claim 2; Page 58; 67pp; French.

Use of the scFvRalpha/gp130/LiRbeta complex, used to modulate induce phosphorylation of the tyrosine of gp130 and LiRbeta, or to cellularly where cells expressing the receptor complex are in the periphery or peripheral nervous system, in neurons implicated in neuro-muscular function or in skeletal muscle. The complex or antibodies are also used to decrease the survival, growth or proliferation of tumour cells or to facilitate the proliferation and/or inhibit differentiation of cells stocks. The complex is also used to modulate activity of the gp130/LiRbeta receptor or cells expressing that receptor, particularly those cells implicated in the immune, haematopoietic, nervous or reproductive system, the liver or skeletal muscle. Molecules of the invention may be used to prevent or treat neurodegenerative diseases including amyotrophic lateral sclerosis, Parkinson's and Huntington's disease, to repair or regenerate nervous

CC or muscular tissue or to maintain muscular mass in paralytic patients.  
CC They may also be used to treat cancer, obesity and associated diseases  
CC and to improve fertility, particularly to avoid endometriosis and/or  
CC assist hysteroscopic implantation, thrombosis, or retinal disease,  
CC particular retinal pigmentosis.  
XX  
Sequence 225 AA;  
50

Query Match	100.0%;	Score 1226;	DB 22,	Length 225
Best Local Similarity	100.0%;	Pred. No. 5.6e-120.		
Matches 225;	Conserved			

	QY	0; Gaps	0
1	MDLRAGDSWGMACICTVLMHLPAYPALNRTGDPGPSPISQKFTVDLIKYLEHGLAKSLAGT	60	
b			
1	mdlragswgmactictvlmhlpaypalnrtgdpgpsisqkftvdlikylehglakslagt	60	

61 YLVYIGPPFNEPDDFNPRLEAETLPRAVDLEVRSKLNKDKLELTQNYFAYSHLCYLRGCT 60

121 NROATAE LRSLAHCTSLGILGSIAGVAAALGYPLDPLPSTELPMWDCRAUWCTA  
122

181 KMDFWLLKELQTLILWSAKDNRLKKKMQPPAAAVTTHIGARCE 207  
|||||.....  
.....cslgllglsglmaalgyrlpprlpqteptwlpqphsdflq 180

RESULT 6

AAW25831 standard; Protein; 253 AA.

16-OCT-2001 (first entry

Human; cancer; ulcer; HIV infection; antiinflammatory agent; sequence SEQ ID NO:1346

[illegible][illegible][illegible]

neurodegenerative disorder; Parkinson's disease; depression; neurological disorder.

W0200153455-A2  
26-TUR-2001

22-DEC-2000; 2000WO-US35017

21-JAN-2000; 2000US-0488725.  
25-APR-2000; 2000US-0552317.

Tang YT, Liu C, Drmanac RT

WPI; 2001-457603/49.



KW hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;  
 KW therapy.  
 XX  
 OS Mus sp.

Key Location/Qualifiers  
 FT Peptide 1..27  
 FT /label= Sig-peptide  
 FT 28..225  
 FT Protein /label= Mat.protein

PN MO9833922-A1.

XX 06-AUG-1998.

XX 02-FEB-1998; 98WO-US02363.

XX 30-JAN-1998; 98US-0016534.

XX 03-FEB-1997; 97US-0792019.

XX (AMGE-) AMGEN INC.

XX Chang M, Elliot GS, Sarmiento U, Senaldi G.

XX WPI: 1998-437475/37.

XX N-PSDB; AAV47512.

PS Claim 13; Fig 5; 120pp; English.  
 XX This is the amino acid sequence of a murine neurotrophic factor,  
 CC designated NNT-1, that is a growth factor for neurons and for B or  
 CC T cells. It was deduced from isolated NNT-1 cDNA (see AAV47512).  
 CC Human NNT-1 (see AAM29715) is also provided. Vectors and host cells  
 CC for use in the production of human murine recombinant NNT-1  
 CC polypeptides. These are used to treat: (i) neurological or  
 CC immunological diseases, specifically Alzheimer's, Parkinson's  
 CC or Huntington's diseases, amyotrophic lateral sclerosis,  
 CC Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy and  
 CC degeneration of the neural retina, or conditions characterised by T  
 CC or B cell defects, e.g. common variable immunodeficiency (CVID),  
 CC selective IGA deficiency, hypogammaglobulinaemia and X-linked  
 CC agammaglobulinaemia (XLA), but many others disclosed; and (ii)  
 CC inflammation. NNT-1 is also able to boost immunoreactivity and  
 CC antibody production following vaccination, and, since it inhibits  
 CC tumour necrosis factor production, it may also be useful for  
 CC treating sepsis. In addition, cells that have been engineered to  
 CC express NNT-1 can be implanted, or nucleic acids are delivered in  
 CC gene therapy vectors.  
 CC  
 XX

Sequence 225 AA;

Query Match 97.3%; Score 1193; DB 19; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 1.6e-116;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

XX 1 MDLRAGSGMGLACLCVLMHLPALNRTGDPGPGPSIQKTYDLTRYLEHQLSLAGT 60  
 Db 1 mdlttagsgwmglactclvhlmpalnrctgdpqpsiqktydltrylehgqlslagt 60  
 YY 61 YLNTLGPFPNEPDPNPRIGAEFLPRATVDLEWRSINDKRLKLTQNYEAYSHLLCYRGL 120  
 Db 61 ylnylgppfnepdpnprigaelpratvdlwrsindkrlrltqnyeahshllcylrgl 120  
 YY 121 NRQAATRLRLSLAHFCTSLQGLGSIAGVMAALGYPLPQPLPGTEPMTWTPGPAHSDFLQ 180  
 Db 121 nrqaatrlrlslahfctslqglgsiagvmaatlgyp-lpgplpgtepawpapsdfllq 180  
 YY 181 KMDFFWLKELQTWLWRSKAKDENRLKKKQPPAAVATVTLHGAHGF 225

Db 181 kmddfwllkelqtwlwrskakdenrlkkkqppaavatlhgaht 225

RESULT 9  
 AAM56142  
 ID AAM56142 standard; Protein; 225 AA.

XX AAM56142;

XX 13-JUL-1998 (first entry)

XX Amino acid sequence of murine neurotrophic factor NNT-1.

XX Mouse; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;  
 KW treatment; neurological disease; degeneration; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; ALS; Alzheimer's disease; stroke.  
 XX

OS Mus sp.

Key Location/Qualifiers  
 FT Peptide 1..27  
 FT /note= "signal peptide"  
 FT 28..225  
 FT Protein /note= "mature peptide"

XX US5741772-A.

XX 21-APR-1998.

XX 03-FEB-1997; 97US-0792019.

XX 03-FEB-1997; 97US-0792019.

XX (AMGE-) AMGEN INC.

XX Chang M;

XX WPI: 1998-260526/23.

XX N-PSDB; AAV22654.

XX Neurotrophic factor NNT-1 polypeptide and related nucleic acids -  
 PT useful for stimulating growth of motor and sympathetic neurons  
 CC  
 CC Claim 2; Fig 5; 41pp; English.  
 CC  
 CC The present sequence represents a murine neurotrophic factor, designated  
 CC NNT-1, which is capable of stimulating growth of motor or sympathetic  
 CC neurons. The NNT-1 protein is useful in the treatment of neurological  
 CC diseases characterised by the degeneration and death of particular  
 CC classes of neurons. These diseases specifically include Parkinson's  
 CC disease, amyotrophic lateral sclerosis (ALS), Alzheimer's disease,  
 CC stroke and various degenerative disorders affecting vision.  
 CC  
 XX

Sequence 225 AA;

Query Match 97.3%; Score 1193; DB 19; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 1.6e-116;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

XX 1 MDLRAGSGMGLACLCVLMHLPALNRTGDPGPGPSIQKTYDLTRYLEHQLSLAGT 60  
 Db 1 mdlttagsgwmglactclvhlmpalnrctgdpqpsiqktydltrylehgqlslagt 60  
 YY 61 YLNTLGPFPNEPDPNPRIGAEFLPRATVDLEWRSINDKRLKLTQNYEAYSHLLCYRGL 120  
 Db 61 ylnylgppfnepdpnprigaelpratvdlwrsindkrlrltqnyeahshllcylrgl 120  
 YY 121 NRQAATRLRLSLAHFCTSLQGLGSIAGVMAALGYPLPQPLPGTEPMTWTPGPAHSDFLQ 180  
 Db 121 nrqaatrlrlslahfctslqglgsiagvmaatlgyp-lpgplpgtepawpapsdfllq 180  
 YY 181 KMDFFWLKELQTWLWRSKAKDENRLKKKQPPAAVATVTLHGAHGF 225

OY 181 KMDDFWLLKELQTWLWRSKADFNRLKKKQPPAAVTLHLGHGF 225  
 |||||  
 Db 181 kmddfllkqlgwlsrskdfrlkkkmpaasvllhahgf 225

## RESULT 10

AA87814  
 ID AA87814 standard; Protein; 225 AA.

AA87814;

24-AUG-2000 (first entry)

Murine NNT-1 protein.

NNT-1; neurotrophic factor; nootropic; neuroprotective; treatment; anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological; nervous system degeneration; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome; murine; Huntington's disease; peripheral neuropathy; neural retina degeneration; retinopathy; immune disorder; hematopoietic disorder.

Mus sp.

US6054294-A.

25-APR-2000.

12-DEC-1997; 97US-0988819.

03-FEB-1997; 97US-0792019.

(AMGE-) AMGEN INC.

Chang M;

WPI: 2000-338492/29.

N-PSDB: AAA39483.

New nucleic acids encoding neurotrophic factors useful for stimulating growth of motor or sympathetic neurons for treating neuron cell damage

Claim 2b; Fig 5; 42pp; English.

This invention describes a novel nucleic acid molecule (I) encoding a novel neurotrophic factor (NNT-1) (II) which has nootropic, neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and ophthalmological activity. (I) is useful for producing NNT-1 polypeptides which are useful for treating patients in whom various cells of the central, autonomic, or peripheral nervous system have degenerated and/or have been damaged by congenital disease, trauma, mechanical damage, surgery, stroke, ischemia, infection, metabolic disease, nutritional deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral neuropathy induced by diabetes or other metabolic disorders, and/or dystrophies or degeneration of the neural retina such as retinitis pigmentosa, drug-induced retinopathies, stationary forms of night blindness, progressive cone-rod degeneration, immune disorders and hemopoietic disorders. (I) is effective in treating neurological conditions and promotes neuron regeneration. Neural functions are effectively restored in patients suffering from various neurological disorders. This sequence represents the murine NNT-1 protein described in the method of the invention.

Sequence 225 AA;

Query Match 97.3%; Score 1193; DB 21; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 1,6e-116;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 MDLRAGDSWGMALACCTVYVHLPAVPALNRTGPGGPGSTOKTYDFTVAT LKSLAGT 60  
 |||||  
 Db 1 mdlragdswgmalcctvylwhlpavpalnrtgpgpgpsiqtdydlctydlcslagt 60  
 OY 61 YLNIYIPPPENPPENPPRIGAEFTLPRAVDLEWRSINDKRLTQNEVYSVLLCYLGL 120  
 |||||  
 Db 61 yllylpppennppennpprigracelpratyvlewrsindrlrtgyeayshllylgl 120  
 OY 121 NRQAATAELEKRSIAHCTSLQGLGSLAGYMAALGTPLPQPGREPTWPGPAHSDFEQ 180  
 |||||  
 Db 121 nrgaataelrslahctslqglgslagymaalgyplpqpjgpcwapaapahsdfiq 180  
 OY 181 KMDDFWLLKELQTWLWRSKADFNRLKKKQPPAAVTLHLGHGF 225  
 |||||  
 Db 181 kmddfllkqlgwlsrskdfrlkkkmpaasvllhahgf 225

## RESULT 11

AA819586  
 ID AA819586 standard; Protein; 215 AA.

AA819586;

22-JAN-2001 (first entry)

Human interleukin-B60 (IL-B60).

Interleukin-B60; IL-B60; human; cytokine; cytokine-like factor-1;

haematopoietic; inflammation; antinflammatory; autoimmune disease;

therapy.

Homo sapiens.

Key Location/Qualifiers

peptide 1..17

protein /label= signal\_peptide

18..215

/label= Mature-protein

WO200053631-A1.

14-SEP-2000.

09-MAR-2000; 2000MO-US06182.

11-MAR-1999; 99US-0267901.

(SCHE ) SCHERING CORP.

Oppmann B, Timans JC, Kastelein RA, Bazan JF;

WPI: 2000-587426/55.

N-PSDB: AAA88546.

Cytokine-like factor 1 (CLF-1) and interleukin (IL)-B60 complexes,

polypeptides, and nucleic acids, useful in research, diagnosis and for

treating inflammatory and autoimmune disorders.

Claim 1; Page 15-16; 97pp; English.

The present sequence is that of human interleukin-B60 (IL-B60), a novel, small soluble cytokine-like protein that exhibits structural motifs characteristic of a member of the long-chain cytokines, and which shows homology to granulocyte colony stimulating factor and interleukin-6. IL-B60 may have either stimulatory or inhibitory effects on haematopoietic cells, including e.g. lymphoid cells, such as T-cells, B-cells, natural killer cells, macrophages, dendritic cells, haematopoietic progenitors, etc. Methods are provided for modulating the physiology or development of a cell or tissue culture cells by contacting the cell with an agonist or antagonist of IL-B60 or an agonist of antagonist of a complex of mature IL-B60 and its partner, cytokine-like factor-1 (CLF-1, see AAB19588). The IL-B60/CLF-1 cytokine serves as a key physiological

CC factor in motor neuron development and regeneration. IL-60B, its  
 CC agonists and antagonists may be used to treat inflammatory or  
 CC autoimmune disorders and also for drug screening.  
 XX  
 SQ Sequence 215 AA;

Query Match 95.4%; Score 1169; DB 21; Length 215;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-114;  
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 MLCALCTVLMHLPVAPVNLNRTGDPGPGSPSTQKTYDTRILEHOLSLACTYLYNVLGPPN 70  
 Db 1 mlaclctvlwhlpavpalnrtgdpgpspsiktydltrylehqrlsagtylnylgppn 60  
 QY 71 EPPFNPPRIGAEETLPRAVTDLEWRSINDKRLTQWNEAVSHLCTYLRGLNQAAATAEIR 130  
 Db 61 epdlnpprllgaetlprativdlewrsindkrltqweavshlcytlnrqaataelr 120  
 QY 131 RSLAHFCTSLGSLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLOKMDPEWLKE 190  
 Db 121 rslahfctslgslisagvmaalgyplpqplpgteptwtpgpahsdflokmddfwllke 180  
 QY 191 LQTWLMSAKDENRLKKMKQPPAAVTLHLGAHGF 225  
 Db 181 lqtwlmsakdenrlkkmqppaavtlhlghgf 215

## RESULT 12

AAB19587  
 ID AAB19587 standard; Protein; 215 AA.

AC AAB19587;  
 DT 22-JAN-2001 (first entry)

XX Mouse Interleukin-B60 (IL-B60).

DE Interleukin-B60; IL-B60; mouse; cytokine; cytokine-like factor-1;  
 KW hematopoietic; inflammation; antiinflammatory; autoimmune disease;  
 therapy.  
 XX

OS Mus musculus.

Key Location/Qualifiers  
 FT Peptide 1..17  
 FT /label= Signal\_peptide  
 FT 18..215  
 FT Protein /label= Mature-protein

PN W0200053631-A1.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06182.

PR 11-MAR-1999; 99US-0267901.

PA (SCHE ) SCHERING CORP.

PI Oppmann B, Timans JC, Kastelein RA, Bazan JF;

DR WPI: 2000-587426/55.

DR N-PSDB: AAA8547.

XX Cytokine-like factor 1 (CLF-1) and Interleukin (IL)-B60 complexes,  
 PT polypeptides, and nucleic acids, useful in research, diagnosis and for  
 PR treating inflammatory and autoimmune disorders -  
 XX

PS Claim 1; Page 17; 97pp; English.

XX The present sequence is that of mouse interleukin-B60 (IL-B60), a  
 CC novel, small soluble cytokine-like protein that exhibits structural

CC motifs characteristic of a member of the long-chain cytokines, and  
 CC which shows homology to granulocyte colony stimulating factor and  
 CC interleukin-6. IL-60B may have either stimulatory or inhibitory  
 CC effects on hematopoietic cells, including e.g. lymphoid cells,  
 CC such as T-cells, B-cells, natural killer cells, macrophages,  
 CC dendritic cells, hematopoietic progenitors, etc. Methods are  
 CC provided for modulating the physiology or development of a cell or  
 CC tissue culture cells by contacting the cell with an agonist or  
 CC antagonist of IL-B60 or an agonist of antagonist of a cell or  
 CC mature IL-B60 and its partner, cytokine-like factor-1 (CLF-1, see  
 CC AAB19587). The IL-B60/CLF-1 cytokine serves as a key physiological  
 CC factor in motor neuron development and regeneration. IL-60B, its  
 CC agonists and antagonists may be used to treat inflammatory or  
 CC autoimmune disorders and also for drug screening.  
 XX

SQ Sequence 215 AA;

Query Match 92.7%; Score 1136; DB 21; Length 115;  
 Best Local Similarity 96.7%; Pred. No. 1.3e-110;  
 Matches 208; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 11 MLCALCTVLMHLPVAPVNLNRTGDPGPGSPSTQKTYDTRILEHOLSLACTYLYNVLGPPN 70  
 Db 1 mlaclctvlwhlpavpalnrtgdpgpspsiktydltrylehqrlsagtylnylgppn 60  
 QY 71 EPPFNPPRIGAEETLPRAVTDLEWRSINDKRLTQWNEAVSHLCTYLRGLNQAAATAEIR 130  
 Db 61 epdlnpprllgaetlprativdlewrsindkrltqweavshlcytlnrqaataelr 120  
 QY 131 RSLAHFCTSLGSLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLOKMDPEWLKE 190  
 Db 121 rslahfctslgslisagvmaalgyplpqplpgteptwtpgpahsdflokmddfwllke 180  
 QY 191 LQTWLMSAKDENRLKKMKQPPAAVTLHLGAHGF 225  
 Db 181 lqtwlmsakdenrlkkmqppaavtlhlghgf 215

## RESULT 13

AAM20115  
 ID AAM20115 standard; Protein; 164 AA.

AC AAM20115;

DT 12-OCT-2001 (first entry)

DE Peptide #6549 encoded by probe for measuring cervical gene expression.  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX

OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-488901/53.  
 DR

XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27: SEQ ID No 24941; 487bp; English.  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see A110068-A118459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 164 AA:  
 SQ

Query Match 72.2%; Score 885; DB 22; Length 164;  
 Best Local Similarity 99.4%; Pred. No. 1.6e-84;  
 Matches 163; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 LNYGPPFNEPFPRLGAETLPRAVDLEWRSINDKRLTONVEAYSHLCYLRGLN 121  
 Db 1 lnylppfnepfpnprlgaetlpratvdlewrsindkrltqnyeyshlcylyrqln 60  
 QY 122 RQATTAELRRSLAHFCTSLUGLIGSIAGVMAALGYPPLPQPLPTEPTWTPGPAHSDFLQK 181  
 Db 61 rqaataelrrslahfctsluglgsiagvmaalyplpqpplptepwtpgpahsdflqk 120  
 QY 182 MDQFWLKELOTWLMRSADKDFNRKKMQPPAAVTLHGAGHF 225  
 Db 121 mdfwllkelqtwlmsakdfnrkkmqppaavtlhlgahgf 164

RESULT 14  
 AAM34012  
 ID AAM34012 standard; Protein: 164 AA.

XX AAM34012;  
 AC  
 XX 17-OCT-2001 (first entry)  
 DT  
 XX Peptide #8049 encoded by probe for measuring placental gene expression.  
 DE  
 XX Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 KM  
 XX Homo sapiens.  
 OS  
 XX WO200157272-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00663.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0633366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PR  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488897/53.  
 XX

PT Human genome-derived single exon nucleic acid probes use ul for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 27: SEQ ID No 34281; 654bp; English.  
 CC The present invention relates to single exon nucleic acid probes (SENP:  
 CC see A1131315-A1157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 CC  
 CC Sequence 164 AA:  
 SQ

Query Match 72.2%; Score 885; DB 22; Length 164;  
 Best Local Similarity 99.4%; Pred. No. 1.6e-84;  
 Matches 163; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 LNYGPPFNEPFPRLGAETLPRAVDLEWRSINDKRLTONVEAYSHLCYLRGLN 121  
 Db 1 lnylppfnepfpnprlgaetlpratvdlewrsindkrltqnyeyshlcylyrqln 60  
 QY 122 RQATTAELRRSLAHFCTSLUGLIGSIAGVMAALGYPPLPQPLPTEPTWTPGPAHSDFLQK 181  
 Db 61 rqaataelrrslahfctsluglgsiagvmaalyplpqpplptepwtpgpahsdflqk 120  
 QY 182 MDQFWLKELOTWLMRSADKDFNRKKMQPPAAVTLHGAGHF 225  
 Db 121 mdfwllkelqtwlmsakdfnrkkmqppaavtlhlgahgf 164

RESULT 15  
 AAY09197  
 ID AAY09197 standard; Protein: 208 AA.

XX AAY09197;  
 AC  
 XX 20-JUL-1999 (first entry)  
 DT  
 XX Human DMAX Interleukin-40 (DIL-40) alternative sequence.  
 DE  
 XX DMAX Interleukin-40; DIL-40; activation regulation; development;  
 KW differentiation; hematopoietic; neural cell; drug target; drug;  
 KM immunological condition.  
 XX  
 OS Homo sapiens.  
 XX WO9919491-A2.  
 PN  
 XX 22-APR-1999.  
 PD  
 XX 13-OCT-1998; 98WO-US20932.  
 PF  
 XX 14-OCT-1997; 97US-0950238.  
 PR  
 XX (SCHE ) SCHERING CORP.  
 PA Bazan JF, Kastelein RA;  
 PI WPI; 1999-312475/26.  
 DR N-PSDB; AAX60797.  
 DR  
 XX New human DMAX Interleukin-40 (DIL-40) polypeptide and antagonists  
 PT Claim 1; Page 10-11; 76pp; English.  
 PS  
 XX The invention relates to a human DMAX interleukin-40 (DIL-40) protein.  
 CC Host cells containing a vector comprising the DIL-40 nucleic acid can be  
 CC used for the recombinant expression of the protein. The DIL-40 protein,  
 CC gene and antibody are useful for regulating activation, development  
 CC differentiation and function of various cell types, including  
 CC hematopoietic or neural cells. The antibodies may be antagonists, and





Mon Mar 18 11:23:39 2002

us-09-931-704-2.ra1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:28:54 ; Search time 30.45 seconds  
(without alignments)  
166,280 Million cell updates/sec

Title: US-09-931-704-2  
Perfect score: 1226  
Sequence: 1 MDLRAGDSWGMALACTIVLM.....KKKKPPAAAVTLHGANGF 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues  
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1226	100.0	225	US-08-792-019B-2	Sequence 2, Appli
2	1226	100.0	225	US-09-106-182-2	Sequence 2, Appli
3	1226	100.0	225	US-08-988-819-2	Sequence 2, Appli
4	1226	100.0	225	US-09-016-534-2	Sequence 2, Appli
5	1193	97.3	225	US-08-792-019B-5	Sequence 5, Appli
6	1193	97.3	225	US-08-988-819-5	Sequence 5, Appli
7	1193	97.3	225	US-09-016-534-5	Sequence 5, Appli
8	124.5	10.2	203	US-09-106-182-3	Sequence 3, Appli
9	118.5	9.7	203	US-08-233-609-3	Sequence 3, Appli
10	118.5	9.7	203	US-08-444-083-3	Sequence 3, Appli
11	118.5	9.7	203	US-08-286-304-3	Sequence 3, Appli
12	118.5	9.7	203	US-08-442-745-3	Sequence 3, Appli
13	118.5	9.7	203	US-08-443-129-3	Sequence 3, Appli
14	118.5	9.7	203	US-08-443-952-3	Sequence 3, Appli
15	118.5	9.7	203	US-08-443-130-3	Sequence 3, Appli
16	118.5	9.7	203	US-08-898-911-3	Sequence 3, Appli
17	118.5	9.7	203	US-08-898-911-3	Sequence 3, Appli
18	118.5	9.7	203	US-08-444-083-8	Sequence 8, Appli
19	118.5	9.7	203	US-08-444-083-8	Sequence 8, Appli
20	118.5	9.7	203	US-08-442-745-8	Sequence 8, Appli
21	118.5	9.7	203	US-08-443-129-8	Sequence 8, Appli
22	118.5	9.7	203	US-08-443-952-8	Sequence 8, Appli
23	118.5	9.7	203	US-08-443-130-8	Sequence 8, Appli
24	118.5	9.7	203	US-08-898-911-8	Sequence 8, Appli
25	118.5	9.7	203	US-08-898-911-8	Sequence 8, Appli
26	118.5	9.7	203	US-08-898-911-8	Sequence 8, Appli
27	118.5	9.7	203	US-08-898-911-8	Sequence 8, Appli

28	91.5	7.5	201	4	US-09-016-534-11	Sequence 11, Appli
29	91.5	7.5	195	1	PCT-US95-04467-8	Sequence 8, Appli
30	91	7.4	195	1	US-07-959-284-5	Sequence 5, Appli
31	91	7.4	195	2	US-08-308-736A-5	Sequence 2, Appli
32	91	7.4	195	5	PCT-US92-08258-2	Sequence 2, Appli
33	91	7.4	195	5	PCT-US93-09649A-5	Sequence 5, Appli
34	91	7.4	195	5	PCT-US93-09649-5	Sequence 5, Appli
35	85.5	6.9	200	4	US-08-756-317-11	Sequence 11, Appli
36	84	6.7	200	4	US-08-949-155-4	Sequence 4, Appli
37	82.5	6.7	582	2	US-08-422-705B-9	Sequence 9, Appli
38	82.5	6.7	582	2	US-07-959-284-8	Sequence 8, Appli
39	82	6.7	200	1	US-07-959-284-8	Sequence 8, Appli
40	82	6.7	200	2	US-08-308-736A-8	Sequence 8, Appli
41	82	6.7	200	2	PCT-US93-09649A-8	Sequence 8, Appli
42	82	6.7	200	5	PCT-US93-09649-8	Sequence 8, Appli
43	81	6.6	200	1	US-07-959-284-6	Sequence 6, Appli
44	81	6.6	200	2	US-08-308-736A-6	Sequence 6, Appli
45	81	6.6	200	5	PCT-US93-09649A-6	Sequence 6, Appli

## ALIGNMENTS

RESULT 1  
US-08-792-019B-2  
Sequence 2, Appli Application US/08792019B  
Patent No. 5741772  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: THE NEUROTROPIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: 1840 DEHAVILLAND DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/792,019B  
FILING DATE: 03-FEB-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442  
INFORMATION FOR SEQ. ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-792-019B-2  
Query Match  
Best local similarity 100.0%; Score 1226; DB 1; Length 225;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 NRQAATAELRSLAHFCTSLQGLLSIAGVMAALGYPLPQPLPCTEPTWTGPAHSDFLQ 180  
QY 181 KMDDFWLKELQOTWLMRSKADFNRLKKMKOPPAAVTLHGAHGF 225  
Db 181 KMDDFWLKELQOTWLMRSKADFNRLKKMKOPPAAVTLHGAHGF 225

## RESULT 2

US-09-106-182-2  
Sequence 2, Application US/09106182  
Patent No. 6046035  
GENERAL INFORMATION:  
APPLICANT: Shi, Yangu  
APPLICANT: Ruben, Steve  
TITLE OF INVENTION: Cardiotrophin-Like Cytokine  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc  
STREET: 9410 Key West Ave  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,182  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/051,053  
ATTORNEY/AGENT INFORMATION:  
FILING DATE: 30-JUN-1997  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF385  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-106-182-2

Query Match 100.0%; Score 1226; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3, 1e-129;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLRAGDSWGMACCTVLMHLPVAVPALNRTGDPGPGPSIOKTYDLTRYLEHQLRSLAGT 60  
Db 1 MDLRAGDSWGMACCTVLMHLPVAVPALNRTGDPGPGPSIOKTYDLTRYLEHQLRSLAGT 60  
QY 61 YLNYIGPPFNEPDNPRLGAETLPRAIVDLEWWSLNDKRLTONYEAYSHLLCYLGL 120  
Db 61 YLNYIGPPFNEPDNPRLGAETLPRAIVDLEWWSLNDKRLTONYEAYSHLLCYLGL 120  
QY 121 NRQAATAELRSLAHFCTSLQGLLSIAGVMAALGYPLPQPLPCTEPTWTGPAHSDFLQ 180  
Db 121 NRQAATAELRSLAHFCTSLQGLLSIAGVMAALGYPLPQPLPCTEPTWTGPAHSDFLQ 180  
QY 181 KMDDFWLKELQOTWLMRSKADFNRLKKMKOPPAAVTLHGAHGF 225  
Db 181 KMDDFWLKELQOTWLMRSKADFNRLKKMKOPPAAVTLHGAHGF 225

## RESULT 3

US-08-988-819-2  
Sequence 2, Application US/08988819  
Patent No. 6054294  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,819  
FILING DATE: 12-DEC-1997  
CLASSIFICATION: 536  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
ATTORNEY/AGENT INFORMATION:  
FILING DATE: 03-FEB-1997  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442A  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-988-819-2

Query Match 100.0%; Score 1226; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3, 1e-129;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLRAGDSWGMACCTVLMHLPVAVPALNRTGDPGPGPSIOKTYDLTRYLEHQLRSLAGT 60  
Db 1 MDLRAGDSWGMACCTVLMHLPVAVPALNRTGDPGPGPSIOKTYDLTRYLEHQLRSLAGT 60  
QY 61 YLNYIGPPFNEPDNPRLGAETLPRAIVDLEWWSLNDKRLTONYEAYSHLLCYLGL 120  
Db 61 YLNYIGPPFNEPDNPRLGAETLPRAIVDLEWWSLNDKRLTONYEAYSHLLCYLGL 120  
QY 121 NRQAATAELRSLAHFCTSLQGLLSIAGVMAALGYPLPQPLPCTEPTWTGPAHSDFLQ 180  
Db 121 NRQAATAELRSLAHFCTSLQGLLSIAGVMAALGYPLPQPLPCTEPTWTGPAHSDFLQ 180  
QY 181 KMDDFWLKELQOTWLMRSKADFNRLKKMKOPPAAVTLHGAHGF 225  
Db 181 KMDDFWLKELQOTWLMRSKADFNRLKKMKOPPAAVTLHGAHGF 225  
RESULT 4  
US-09-016-534-2  
Sequence 2, Application US/09016534  
Patent No. 6143874  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
APPLICANT: ELLIOTT, GARY S.  
APPLICANT: SARIENTO, ULLA  
TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:

Mon Mar 18 11:23:39 2002

us-09-931-704-2.ral

Page 3

ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,534  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442B  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-016-534-2

Query Match 100.0%; Score 1226; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3, 1e-129; Indels 0; Gaps 0;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLRAGDSMGLACLTCTVIMHLPAYPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
DB 1 MDLRAGDSMGLACLTCTVIMHLPAYPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
QY 61 YLNTLGPPFNEPDPNPRIGAEITLPRATVDLEWRSNDKRLTONYEAVSHLLCYLRGL 120  
DB 61 YLNTLGPPFNEPDPNPRIGAEITLPRATVDLEWRSNDKRLTONYEAVSHLLCYLRGL 120  
QY 121 NROAATAEELRSLAHFCTSLQGLGSIAGYMAALGYPLPQPLPCTEPTWTPGPAHSDFLQ 180  
DB 121 NROAATAEELRSLAHFCTSLQGLGSIAGYMAALGYPLPQPLPCTEPTWTPGPAHSDFLQ 180  
QY 181 KMDDFWLKELQTWLMRSARKDFNLRKKMOPPAAVTLHLGAHGF 225  
DB 181 KMDDFWLKELQTWLMRSARKDFNLRKKMOPPAAVTLHLGAHGF 225

RESULT 5  
US-08-792-019B-5  
Sequence 5, Application US/08792019B  
Patent No. 5741772  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: 1840 DEHAVILLAND DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/792,019B  
FILING DATE: 03-FEB-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-792-019B-5

Query Match 97.3%; Score 1193; DB 1; Length 225;  
Best Local Similarity 96.9%; Pred. No. 1, 5e-125; Indels 4; Gaps 0;  
Matches 218; Conservative 3; Mismatches 4; Indels 4; Gaps 0;

QY 1 MDLRAGDSMGLACLTCTVIMHLPAYPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
DB 1 MDLRAGDSMGLACLTCTVIMHLPAYPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
QY 61 YLNTLGPPFNEPDPNPRIGAEITLPRATVDLEWRSNDKRLTONYEAVSHLLCYLRGL 120  
DB 61 YLNTLGPPFNEPDPNPRIGAEITLPRATVDLEWRSNDKRLTONYEAVSHLLCYLRGL 120  
QY 121 NROAATAEELRSLAHFCTSLQGLGSIAGYMAALGYPLPQPLPCTEPTWTPGPAHSDFLQ 180  
DB 121 NROAATAEELRSLAHFCTSLQGLGSIAGYMAALGYPLPQPLPCTEPTWTPGPAHSDFLQ 180  
QY 181 KMDDFWLKELQTWLMRSARKDFNLRKKMOPPAAVTLHLGAHGF 225  
DB 181 KMDDFWLKELQTWLMRSARKDFNLRKKMOPPAAVTLHLGAHGF 225

RESULT 6  
US-08-988-819-5  
Sequence 5, Application US/08988819  
Patent No. 6054294  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,819  
FILING DATE: 12-DEC-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442A  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-988-819-5

Query Match  
Best Local Similarity 97.3%; Score 1193; DB 3; Length 225;  
Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMACCTVLMHLPAPVPAALNRGDPGPGSIKTYDRLRYLEHQLRSLAGT 60  
DB 1 MDLRAGDSWGMACCTVLMHLPAPVPAALNRGDPGPGSIKTYDRLRYLEHQLRSLAGT 60  
QY 61 YLNYIGPPNEDPDPRLGATLPRATVDLEWRSNDKRLQNEAVSHLLCYRGL 120  
DB 61 YLNYIGPPNEDPDPRLGATLPRATVDLEWRSNDKRLQNEAVSHLLCYRGL 120  
QY 121 NROAATAELRSLAHFCTSLQGLGSLAGVMAALGYPLPQPLGTEPTWTTPGASHDFLQ 180  
DB 121 NROAATAELRSLAHFCTSLQGLGSLAGVMAALGYPLPQPLGTEPTWTTPGASHDFLQ 180  
QY 181 KMDDFWLKELQTYLWMSAKDENLKKMKQPPAAVTLHGANGF 225  
DB 181 KMDDFWLKELQTYLWMSAKDENLKKMKQPPAAVTLHGANGF 225

## RESULT 7

US-09-016-534-5  
Sequence 5, Application US/09016534  
Patent No. 6143874  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
APPLICANT: ELLIOTT, GARY S.  
APPLICANT: SAMIENTO, ULLA  
TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,534  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442B  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-016-534-5

Query Match  
Best Local Similarity 97.3%; Score 1193; DB 4; Length 225;  
Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MDLRAGDSWGMACCTVLMHLPAPVPAALNRGDPGPGSIKTYDRLRYLEHQLRSLAGT 60

DB 1 MDLRAGDSWGMACCTVLMHLPAPVPAALNRGDPGPGSIKTYDRLRYLEHQLRSLAGT 60  
QY 61 YLNYIGPPNEDPDPRLGATLPRATVDLEWRSNDKRLQNEAVSHLLCYRGL 120  
DB 61 YLNYIGPPNEDPDPRLGATLPRATVDLEWRSNDKRLQNEAVSHLLCYRGL 120  
QY 121 NROAATAELRSLAHFCTSLQGLGSLAGVMAALGYPLPQPLGTEPTWTTPGASHDFLQ 180  
DB 121 NROAATAELRSLAHFCTSLQGLGSLAGVMAALGYPLPQPLGTEPTWTTPGASHDFLQ 180  
QY 181 KMDDFWLKELQTYLWMSAKDENLKKMKQPPAAVTLHGANGF 225  
DB 181 KMDDFWLKELQTYLWMSAKDENLKKMKQPPAAVTLHGANGF 225

## RESULT 8

US-09-106-182-3  
Sequence 3, Application US/09106182  
Patent No. 6046035  
GENERAL INFORMATION:  
APPLICANT: Shi, Yangu  
APPLICANT: Ruben, Steve  
TITLE OF INVENTION: Cardiotrophin-Like Cytokine  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc  
STREET: 9410 Key West Ave  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,182  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/051,053  
FILING DATE: 30-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF385  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 203 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-106-182-3

Query Match  
Best Local Similarity 10.2%; Score 124.5; DB 3; Length 203;  
Matches 50; Conservative 24; Mismatches 85; Indels 19; Gaps 5;

QY 40 IOKTYDRLRYLEHQLRSLAGTLYLVNLYGPPNEDPDPRLGATLPRATVDLEWRS 96  
DB 27 IROTHNLARLLTKYVADQLELYVOGGEPPGCLGFSPPRLPLAGSGPAPSHAGLPV--- 83  
QY 97 LNDKRLQNEAVSHLLCYRGLNROA-----TAEIRSLAHFCTSLQGLGSLAGVMA 151  
DB 84 ---SERLRDAALSLALPALLDVRRQAEINBRAPRLRLSLSDADARQVRAAGAAVEYVL 140





APPLICANT: Baker, Joffree  
APPLICANT: Chien, Kenneth  
APPLICANT: King, Kathleen  
APPLICANT: Pennica, Diane  
APPLICANT: Wood, William  
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/443,130  
FILING DATE: 17-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/233609  
FILING DATE: 25-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/286304





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2002, 09:33:29 ; Search time 1350.94 Seconds

(without alignments)  
10001.325 Million cell updates/sec

Title: US-09-931-704-4

Perfect score: 819  
Sequence: 1 tatattaaagcttcgcgcg.....agccacagctcagctgtcgtt 819

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_om: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pl: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_sy: \*  
28: em\_un: \*  
29: em\_vi: \*  
30: em\_htgo\_hum: \*  
31: em\_htgo\_inv: \*  
32: em\_htgo\_rod: \*  
33: em\_htg\_hum: \*  
34: em\_htg\_inv: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_other: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	819	100.0	819	6	AR002597
2	819	100.0	819	10	AF176913
3	669.4	81.7	797	6	AR002595
4	669.4	81.7	797	9	AF176911
5	627	76.6	1689	9	AF172854
6	429.6	52.5	5087	6	AR002596
7	429.6	52.5	5087	6	AF176912
8	429.6	52.5	135116	2	AP002437
9	429.6	52.5	151441	2	AP003419
10	429.6	52.5	169144	2	AC005849
11	46.2	5.6	47958	2	AC091104
12	44.4	5.4	187727	2	AC021142
13	43.2	5.3	6803	9	HS868758
14	42.4	5.2	1357	9	AF035771
15	42.4	5.2	1524	9	AF004900
16	42.4	5.2	1578	9	HS082108
17	42.4	5.2	1600	9	HSFKAI1MR
18	42.4	5.2	1642	6	AR070449
19	42.2	5.2	1901	5	GGTAMBI
20	42.2	5.1	63082	2	AC022663
21	40.8	5.0	123554	9	AB023049
22	40.8	5.0	200000	9	AP000512
23	40.6	5.0	114653	2	AC092411
24	40.4	4.9	6955	1	TFHIIYB
25	40.4	4.9	188947	2	AC087455
26	40.2	4.9	167693	2	AC011639
27	40.2	4.9	180626	2	AC022824
28	40.2	4.9	184485	2	AC009679
29	40.2	4.9	193263	2	AC090971
30	40.2	4.9	1246	6	AX164174
31	39.2	4.8	105935	9	AL355802
32	39.2	4.8	170530	2	AL359977
33	39.2	4.8	192634	2	AC069466
34	38.8	4.7	160312	2	AP003500
35	38.8	4.7	161286	2	AC025120
36	38.6	4.7	70452	9	HS7854E16
37	38.6	4.7	227531	2	AC015705
38	38.4	4.7	16124	1	AX024384
39	38.4	4.7	16124	6	AX024277
40	38.4	4.7	23404	1	SCI0A9
41	38.4	4.7	68750	1	AF210843
42	38.4	4.7	178597	2	AC063962
43	38.2	4.7	1970	5	AB001742
44	38.2	4.7	37770	9	AC010513
45	38.2	4.7	145133	2	AC013822

# ALIGNMENTS

RESULT 1  
AR002597 LOCUS AR002597 819 bp DNA  
SEQUENCE 4 from patent US 5741772.  
AR002597  
VERSION AR002597.1 GI:3964151  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 819)  
AUTHORS Chang, M.  
TITLES Neurotrophic factor NNT-1  
JOURNAL Patent: US 5741772-A 4 21-APR-1998;  
FEATURES location/Qualifiers  
source 1. 819  
BASE COUNT 156 a 288 c 218 g 157 t  
ORIGIN

PAT 1-DEC-1998



```

QY 541 tgcataggagagcttggtactaccactgcccagcctctgcagggagctgagccagctg 600
DB 541 tgcataggagagcttggtactaccactgcccagcctctgcagggagctgagccagctg 600
QY 601 ggcacctgagccctgagccagctgactccctcagagaatgagatgactctgctgtga 660
DB 601 ggcacctgagccctgagccagctgactccctcagagaatgagatgactctgctgtga 660
QY 661 ggaactgagagctggtctatgagctcagccaggaactcgaacggctgtaagaagaat 720
DB 661 ggaactgagagctggtctatgagctcagccaggaactcgaacggctgtaagaagaat 720
QY 721 ggaactgagagctggtctatgagctcagccaggaactcgaacggctgtaagaagaat 780
DB 721 ggaactgagagctggtctatgagctcagccaggaactcgaacggctgtaagaagaat 780
QY 781 ccttaacccccacacactcagagccagctgagctgctt 819
DB 781 ccttaacccccacacactcagagccagctgagctgctt 819

RESULT 3
AR002595 797 bp DNA PAT 04-DEC-1998
LOCUS Sequence 1 from patent US 5741772.
DEFINITION AR002595
ACCESSION AR002595
VERSION AR002595.1 GI:3964149
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 797)
AUTHORS Chang,M.
TITLE Neurotrophic factor NNT-1
JOURNAL Patent: US 5741772-A 1 21-APR-1998;
FEATURES
source 1.797
location/Qualifiers
BASE COUNT 139 a 297 c 218 g 143 t
ORIGIN

Query Match 81.7%; Score 669.4; DB 6; Length 797;
Best Local Similarity 92.0%; Pred. No. 6e-143;
Matches 717; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

```

```

DB 360 ttgagagtggtggcagagcctcgaatgacaaactggcggtgaccaccagaaactgagccctac 419
QY 425 agtaacctctgtgtactgtgagctcaacagctcaagctgcccagctgaaactcga 484
DB 425 agtaacctctgtgtactgtgagctcaacagctcaagctgcccagctgaaactcga 484
QY 485 gtaagctgagccactctgtactcagagctcagagagcctgctgagcagatgagtg 544
DB 485 gtaagctgagccactctgtactcagagctcagagagcctgctgagcagatgagtg 544
QY 545 atgagagagctgtgtactcagagctcagagagcctgctgagcagatgagtgagc 604
DB 545 atgagagagctgtgtactcagagctcagagagcctgctgagcagatgagtgagc 604
QY 605 cctgagcctgagcagagagctctctcagagaatgagtgagctgctgagtgagag 664
DB 605 cctgagcctgagcagagagctctctcagagaatgagtgagctgctgagtgagag 664
QY 665 ctgagagcctgagctgagctgagctcagagagctcagagagctgagtgagtgag 724
DB 665 ctgagagcctgagctgagctgagctcagagagctcagagagctgagtgagtgag 724
QY 725 cctcagagagctcagagagctcagagagctgagtgagtgagtgagtgagtgag 783
DB 725 cctcagagagctcagagagctcagagagctgagtgagtgagtgagtgagtgag 783

RESULT 4
AF176911 797 bp mRNA PRI 04-OCT-1999
LOCUS Homo sapiens neurotrophin-1/B-cell stimulating factor-3 mRNA,
DEFINITION complete cds.
ACCESSION AF176911
VERSION AF176911.1 GI:6007640
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 797)
AUTHORS Senaldi,G., Varnum,B., Sarmiento,U., Lile,J.,
Scully,S., Guo,J., Elliott,G., McNinch,J., Shaklee,C.,
Freeman,D., Mann,F., Simonet,M.S., Boone,T., and Chang,M.S.
TITLE Novel neurotrophin-1/B cell stimulating factor-3: a cytokine of the
IL-6 family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11458-1146. (1999)
MEDLINE 99432254
REFERENCE 2 (bases 1 to 797)
AUTHORS Senaldi,G., Varnum,B., Sarmiento,U., Lile,J.,
Scully,S., Guo,J., Elliott,G., McNinch,J., Shaklee,C.,
Mann,F., Simonet,M.S., Boone,T., and Chang,M.S.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1999) Amgen, Inc., One Amgen Center Drive,
Thousand Oaks, CA 91320, USA
FEATURES
source 1.797
location/Qualifiers
BASE COUNT 139 a 297 c 218 g 143 t
ORIGIN

Query Match 81.7%; Score 669.4; DB 9; Length 797;

```



Db 600 CTGGCTCTGAGAGAGCTGACGACCTGCTGTGGCGCTGCGGCCAAGACCTTCAACGGGCT 659  
 QY 709 taagaagaatgcaagcctcccaagcctcagtcagtcagccttgagagcacttgagagcactggtt 768  
 Db 660 CAAGAAAGAAATGACGCTCCAGAGCTGACGACCTGACCTGACCTGAGGAGGCTCATGGCTT 719  
 QY 769 ctgacccctgacct 783  
 Db 720 CTGACTTCTGACCTT 734

RESULT 6  
 AR002596  
 LOCUS AR002596 5087 bp DNA PAT 04-DEC-1998  
 DEFINITION Sequence 3 from patent US 5741772.  
 ACCESSION AR002596  
 VERSION AR002596.1 GI:3964150  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 1 (bases 1 to 5087)  
 AUTHORS Chang, M.  
 TITLE Neurotrophic factor NNT-1  
 JOURNAL Patent: US 5741772-A 3 21-Apr-1998;  
 FEATURES Location/Qualifiers  
 source 1..5087  
 /organism="unknown"  
 BASE COUNT 992 a 1746 c 1191 g 1158 t  
 ORIGIN

Query Match 52.5%; Score 429.6; DB 6; Length 5087;  
 Best Local Similarity 90.4%; Pred. No. 2.5e-88;  
 Matches 459; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 276 accbtaactactgagggcccttcaacagagcctgaactcaatccctcctgactgagag 335  
 Db 3363 AGCTGAACCTACTGGGCCCCCTTTCAAGCAGCCAGACTTAAACCTCCCGCGCTGGGG 3422  
 QY 336 cagaactctgagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 395  
 Db 3423 CAGAGACTCTGCCAGGCGCCACTGTGACTGTGAGGTGTGCGAAGCCTCAATGACAAAC 3482  
 QY 336 tgcgctgacccagacatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 455  
 Db 3483 TCGGGCTGACCCAGACTGACGAGGCTTACAGCCACTTCTGTGTTACTTGTGGCTCA 3542  
 QY 456 accgctgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 515  
 Db 3543 ACCGCTGAGCTGCGCCTGCTGAGCTGCGCGCAGGCTGCGCCTTGTGACAGCAGCTCC 3602  
 QY 516 agggcctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 575  
 Db 3603 AGGGCTCTGCGGCGCAGCTTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3662  
 QY 576 ctctgcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 635  
 Db 3663 CCGTCTGCTGGAGCTGACCCACTTGGACTCTGCGCTGCGCCTGCGCAGTACTTCTCCGCA 3722  
 QY 636 agatgagactctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 695  
 Db 3722 ACATGAGACACTTTCGGCTGCTGAAGGAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 3782  
 QY 696 acttcaacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 755  
 Db 3783 ACTTCAACCGGCTCAAGAGAGATGACAGCTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3842  
 QY 756 agggcactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 783  
 Db 3843 GGGCTCATGCTTCTGACTTCTGACTT 3870

RESULT 7  
 AF176912  
 LOCUS AF176912 5087 bp DNA PRI 04-DEC-1999  
 DEFINITION Homo sapiens neurotrophin-1/B-cell stimulating factor-3 gene,  
 complete cds.  
 ACCESSION AF176912  
 VERSION AF176912.1 GI:6007642  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 5087)  
 AUTHORS Senaldi, G., Varnum, B.C., Sarmiento, U., Starnes, C., Lille, J.,  
 Scully, S., Guo, J., Elliott, G., McIninch, J., Shalee, C.,  
 Freeman, D., Mann, F., Simonet, S., Boone, T., and Chang, M. S.  
 TITLE Novel neurotrophin-1/B cell-stimulating factor-3: a cytokine of the  
 IL-6 family  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11458-114 (1999)  
 MEDLINE 99432254  
 REFERENCE 2 (bases 1 to 5087)  
 AUTHORS Senaldi, G., Varnum, B., Sarmiento, U., Lille, J., Starnes, C.,  
 Scully, S., Guo, J., Elliott, G., McIninch, J., Freeman, D.,  
 Mann, F., Simonet, S., Boone, T., and Chang, M. S.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-AUG-1999) Amgen, Inc., One Amgen Center Drive,  
 Thousand Oaks, CA 91320, USA  
 FEATURES  
 source  
 1..5087  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 joinr<120..135,1368..1534,3365..3855)  
 /product="neurotrophin-1/B-cell stimulating factor-3"  
 joinr(120..135,1368..1534,3365..3855)  
 /codon\_start=1  
 /product="neurotrophin-1/B-cell stimulating factor-3"  
 /protein\_id="AAE0092.1"  
 /db\_xref="GI:6007643"  
 /translation="MDLRAGDSWMLACTVLMHLPAVPALNRTGDPGPGPSIOKTY  
 DLRYLEHQLRSIAGTYLNLGPPENPDRPRLGAEFLPAVDLWVSLNKLK  
 LTONYEAYSHLICYLRLINROATRELRNSLAHFTCSLOGLSIAGVMAALGYLPQ  
 PLGEIPEWTPPRAHSDFLQKDDDWLKELOTWLMRSKRNLRKKMKOPPAAVTL  
 HIGAHFE"

BASE COUNT 992 a 1746 c 1191 g 1158 t  
 ORIGIN  
 Query Match 52.5%; Score 429.6; DB 9; Length 5087;  
 Best Local Similarity 90.4%; Pred. No. 2.5e-88;  
 Matches 459; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 276 accbtaactactgagggcccttcaacagagcctgaactcaatccctcctgactgagag 335  
 Db 3363 AGCTGAACCTACTGGGCCCCCTTTCAAGCAGCCAGACTTAAACCTCCCGCGCTGGGG 3422  
 QY 336 cagaactctgagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 395  
 Db 3423 CAGAGACTCTGCCAGGCGCCACTGTGACTGTGAGGTGTGCGAAGCCTCAATGACAAAC 3482  
 QY 336 tgcgctgacccagacatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 455  
 Db 3483 TCGGGCTGACCCAGACTGACGAGGCTTACAGCCACTTCTGTGTTACTTGTGGCTCA 3542  
 QY 456 accgctgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 515  
 Db 3543 ACCGCTGAGCTGCGCCTGCTGAGCTGCGCGCAGGCTGCGCCTTGTGACAGCAGCTCC 3602  
 QY 516 agggcctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 575  
 Db 3603 AGGGCTCTGCGGCGCAGCTTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3662  
 QY 576 ctctgcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 635

DB	3663	CGCTGGCTTGGAGACTGAACCACTTGGAGTCTCTGGCCCTCCCACTGACTTCTCTCACA	3722
QY	636	agatgatgactcttcgtgcgtcgtgaagaagactgcagaactgtgctatagcgttcagccaag	695
Db	3723	AATATGACACTTCTGCTGCTGAAGAGACTGCAAGACTCGGCTGTGGCTCTGGCCCAAG	3782
QY	696	acttcaaccggtttaaagaagaatgcagactccagaacttcagtcaccctgcagcttg	755
Db	3783	ACTTCAACCGGCTCAAGAGAGATGCACGCTTCACAGCTGCAGTGCACCTGCAGACTTG	3842
QY	756	aggcacatggtcttcgacctgcacct	783
Db	3843	GGGCTCATGCTTCTGACTTGCACCTT	3870
RESULT	8		
AP002437			
LOCUS	AP002437	135116 bp	DNA
DEFINITION		Homo sapiens chromosome 11 clone RP11-678D20 map 11q13, WORKING	
ACCESSION	AP002437		
VERSION	AP002437.1	GI:8307741	
KEYWORDS	HTG; HTGS_PHASE1; HTGS-DRAFT.		
SOURCE	Homo sapiens DNA, clone:RP11-678D20.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 135116)		
TITLE	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		
JOURNAL	Homo sapiens 135,116 genomic DNA of 11q13		
AUTHORS	2 (bases 1 to 135116)		
	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		
	Direct Submission		
	Submitted (02-JUN-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)		
COMMENT	----- Genome Center		

NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1	7749	contig of	7749	bp in length
7850	15108	contig of	7259	bp in length
15209	23728	contig of	8520	bp in length
23829	29302	contig of	5474	bp in length
29403	36171	contig of	6769	bp in length
36272	42123	contig of	5852	bp in length

42224	48172	contlg of	5549 bp	in length
48273	52700	contlg of	4428 bp	in length
52801	57587	contlg of	4787 bp	in length
57688	61624	contlg of	3937 bp	in length
61725	67756	contlg of	6032 bp	in length
67857	71834	contlg of	3978 bp	in length
71935	76800	contlg of	4666 bp	in length
76901	81069	contlg of	4169 bp	in length
82033	89332	contlg of	4763 bp	in length
86314	86537	contlg of	3224 bp	in length
89638	92572	contlg of	2735 bp	in length
92473	94672	contlg of	2507 bp	in length
95080	96272	contlg of	1193 bp	in length
96373	98109	contlg of	1737 bp	in length
98210	101030	contlg of	2821 bp	in length
101131	103760	contlg of	2630 bp	in length
103861	106330	contlg of	2470 bp	in length
106431	108601	contlg of	2171 bp	in length
108702	110440	contlg of	1739 bp	in length
110541	112280	contlg of	1740 bp	in length
112381	114485	contlg of	2505 bp	in length
114486	117018	contlg of	2033 bp	in length
117119	118839	contlg of	1721 bp	in length
119840	120968	contlg of	2029 bp	in length
121069	122947	contlg of	1879 bp	in length
123048	124231	contlg of	1184 bp	in length
124332	125727	contlg of	1366 bp	in length
125828	127192	contlg of	1365 bp	in length
127293	129019	contlg of	1727 bp	in length
129120	130181	contlg of	1062 bp	in length
130282	131483	contlg of	1202 bp	in length
131584	132595	contlg of	1012 bp	in length
132696	133866	contlg of	1191 bp	in length
133987	135116	contlg of	1130 bp	in length

NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	7749:	contlg of	7749 bp	in length
7750	7849:	gap of	100 bp	
7850	15108:	contlg of	7229 bp	in length
15109	15208:	gap of	100 bp	
15209	23728:	contlg of	8520 bp	in length
23729	23828:	gap of	100 bp	
23829	29302:	contlg of	5474 bp	in length
29303	29402:	gap of	100 bp	
29403	36171:	contlg of	6769 bp	in length
36172	36271:	gap of	100 bp	
36272	42123:	contlg of	5852 bp	in length
42124	42223:	gap of	100 bp	
42224	48172:	contlg of	5949 bp	in length
48173	48272:	gap of	100 bp	
48273	52700:	contlg of	4428 bp	in length
52701	52800:	gap of	100 bp	
52801	57587:	contlg of	4787 bp	in length
57588	57687:	gap of	100 bp	
57688	61624:	contlg of	3937 bp	in length
61625	61724:	gap of	100 bp	
61725	67756:	contlg of	6032 bp	in length
67757	67857:	gap of	100 bp	
67857	71834:	contlg of	3978 bp	in length
71835	71934:	gap of	100 bp	
71935	76800:	contlg of	4666 bp	in length
76801	76900:	gap of	100 bp	
76901	81069:	contlg of	4169 bp	in length
81070	81169:	gap of	100 bp	
81170	81932:	contlg of	766 bp	in length
81933	82032:	gap of	100 bp	

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

*	1	7749:	contig of 7749 bp in length
*	7750	7849:	gap of 100 bp
*	7850	15108:	contig of 7259 bp in length
*	15109	15208:	gap of 100 bp
*	15209	23728:	contig of 8520 bp in length
*	23729	23828:	gap of 100 bp
*	23829	29302:	contig of 5474 bp in length
*	29303	29402:	gap of 100 bp
*	29403	36171:	contig of 6769 bp in length
*	36172	36271:	gap of 100 bp
*	36272	42123:	contig of 5882 bp in length
*	42124	42223:	gap of 100 bp
*	42224	48177:	contig of 5549 bp in length
*	48178	48272:	gap of 100 bp
*	48273	52700:	contig of 4428 bp in length
*	52701	52800:	gap of 100 bp
*	52801	57587:	contig of 4787 bp in length
*	57588	57687:	gap of 100 bp
*	57688	61624:	contig of 3937 bp in length
*	61625	61724:	gap of 100 bp
*	61725	67756:	contig of 6032 bp in length
*	67757	67856:	gap of 100 bp
*	67857	71834:	contig of 3978 bp in length
*	71835	71934:	gap of 100 bp
*	71935	76800:	contig of 4666 bp in length
*	76801	76900:	gap of 100 bp
*	76901	81069:	contig of 4169 bp in length
*	81070	81169:	gap of 100 bp
*	81170	81932:	contig of 763 bp in length
*	81933	82032:	gap of 100 bp

	*	82033	86213:	contig of 4181 bp	in length
	*	86214	86313:	gap of 100 bp	
	*	86314	89537:	contig of 3224 bp	in length
	*	89538	89637:	gap of 100 bp	
	*	89638	92372:	contig of 2735 bp	in length
	*	92373	92472:	gap of 100 bp	
	*	92473	94979:	contig of 2507 bp	in length
	*	94980	95079:	gap of 100 bp	
	*	95080	96272:	contig of 1193 bp	in length
	*	96273	96372:	gap of 100 bp	
	*	96373	98109:	contig of 1737 bp	in length
	*	98110	98209:	gap of 100 bp	
	*	98210	101030:	contig of 2821 bp	in length
	*	101031	101130:	gap of 100 bp	
	*	101131	103760:	contig of 2630 bp	in length
	*	103761	103860:	gap of 100 bp	
	*	103861	106330:	contig of 2470 bp	in length
	*	106331	106430:	gap of 100 bp	
	*	106431	108601:	contig of 2171 bp	in length
	*	108602	108701:	gap of 100 bp	
	*	108702	110440:	contig of 1739 bp	in length
	*	110441	110540:	gap of 100 bp	
	*	110541	112280:	contig of 1740 bp	in length
	*	112281	112380:	gap of 100 bp	
	*	112381	114885:	contig of 2505 bp	in length
	*	114886	114985:	gap of 100 bp	
	*	114986	117018:	contig of 2033 bp	in length
	*	117019	117118:	gap of 100 bp	
	*	117119	118839:	contig of 1721 bp	in length
	*	118840	118939:	gap of 100 bp	
	*	118940	120968:	contig of 2029 bp	in length
	*	120969	121068:	gap of 100 bp	
	*	121069	122947:	contig of 1879 bp	in length
	*	122948	123047:	gap of 100 bp	
	*	123048	124231:	contig of 1184 bp	in length
	*	124232	124331:	gap of 100 bp	
	*	124332	125727:	contig of 1396 bp	in length
	*	125728	125827:	gap of 100 bp	
	*	125828	127192:	contig of 1365 bp	in length
	*	127193	127292:	gap of 100 bp	
	*	127293	129019:	contig of 1727 bp	in length
	*	129020	129119:	gap of 100 bp	
	*	129120	130181:	contig of 1062 bp	in length
	*	130182	130281:	gap of 100 bp	
	*	130282	131483:	contig of 1202 bp	in length
	*	131484	131583:	gap of 100 bp	
	*	131584	132595:	contig of 1012 bp	in length
	*	132596	132695:	gap of 100 bp	
	*	132696	133886:	contig of 1191 bp	in length
	*	133887	133986:	gap of 100 bp	
	*	133987	135116:	contig of 1130 bp	in length.
FEATURES					
source					
			1.	135116	
			/organism="Homo sapiens"		
			/db_xref="taxon:9606"		
			/chromosome="11"		
			/map="11q13"		
			/clone="RP11-678D20"		
misc_feature			1..7749		
			/note="assembly_fragment"		
misc_feature			7850..15108		
			/note="assembly_fragment"		
misc_feature			15209..23728		
			/note="assembly_fragment"		
misc_feature			23829..29302		
			/note="assembly_fragment clone_end:Sp6 vector_side:right"		
misc_feature			29403..36171		
			/note="assembly_fragment"		
misc_feature			36272..42123		
			/note="assembly_fragment"		
misc_feature			42224..48172		
Query Match			52.5%;	Score 429.6;	DB 2; Length 135116;

Matches	459	Conservative	0	Mismatches	49	Indels	0	Gaps	0
Best Local Similarity 90.4%; Pred. No. 1.1e-88;									
Qy	276	acctgaactacttggggccccccttcaagaagcttgactcaatctctcgcgcgtgggg							335
Db	18203	AGCTGAACTACCTGGGCCCCCTTTCAACGAGCCAGCTTCAACCCCTCCGCTGGGG							18262
Qy	336	cagaacctgtgccaggggccaeagtcgaactctgggaagtgtggcaagctctcaatgaagcgc							395
Db	18263	CAGAGACTGTGCCAGGCGCACTGTGACTTGGAGGTTGTGGCAAGCTCAATGACAAAC							18322
Qy	396	tgcgctgacccagaactatgagagcgtacagtcacactcctgtgttaacttgcgttgccctca							455
Db	18323	TGCGGCTGACCCAGAACTACGAGGCGCTACAGCCACTTGTGTACTTGTACTTGGCGGCTCA							18382
Qy	456	accgtgaagctcccaagctgaactccgaagtcgaagtaactcggccactctctaccagcttcc							515
Db	18383	ACCGTAGGCTGCACCTGTGTAGCTGGCGCCGACAGCTGGCCACTTCTGCATCAGCTTCC							18442
Qy	516	aggagctctgtgggcagcatltgcaagtgtcaltgagcgaagcttggcttaccacacttgcacagc							575
Db	18443	AGGGCGTGGGGGACGACATTTGGGGCGTATGAGCGAGCTTGGGCTATCCACTGCCACAG							18502
Qy	576	ctctgccaggagactgagccagcctctggcccttgccctctgcccaagatgaactctctcaca							635
Db	18503	CGCTGGCTGGGACTGAACCCACTTGTGACTCGCGCCCTGCCACAGTACTTCTCTCCACA							18562
Qy	636	agatgatgactctctgctcgtcgaaggagctgcagacactgtgctatgtcgtttttccaag							695
Db	18563	AGATGACACACTTCTGGCTGCTGAGAGAGCTGCAAGACTGGCTGTGGCGTGGCCAAAG							18622
Qy	696	acttcaaccgctcttaagaagaatgtcagcctcgaagcttcagttcaatccctgcgacttg							755
Db	18623	ACTTCAACCGGCTCAAGAAAGATGACAGCTCCAGACGTCGAGTCACTTCCTGACCTGG							18682
Qy	756	aggcacatgtttctgacctgcgacct							783
Db	18683	GGGCTATGGCTTCTGACTTCTGACCTT							18710
RESULT	9								
LOCUS	AP003419/c								
DEFINITION	AP003419 151441 bp DNA HTG 20-MAR-2001								
ACCESSION	AP003419								
VERSION	AP003419.1 GI:13383332								
KEYWORDS	HTG; HTGS; PHASE1; HTGS; DRAFT.								
SOURCE	Homo sapiens DNA, clone:CTD-1337H24.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiinder; Homo.								
TITLE	1 (bases 1 to 151441)								
JOURNAL	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong Seog,P.,								
COMMENT	Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.								
	Direct Submission								
	Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physica								
	and Chemical Research (RIKEN), Genomic Sciences Center (GSC);								
	1-7-22 Suenryo-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan								
	(E-mail:hattori@isc.riken.go.jp, URL:http://hnp.455.riken.go.jp/;								
	TEL:81-45-503-9111, Fax:81-45-503-9170)								
	Genome Center								
	Center: RIKEN Genomic Sciences Center(GSC)								
	Center code: RIKEN								
	Web site: http://hnp.gsc.riken.go.jp/								
	Contact: hattori@isc.riken.go.jp								
	Project Information								
	Center project name: HumDrafl1								
	Center clone name: CTD-1337H24								
	Summary Statistics								
	Sequencing vector: PCR products; 100% of read								
	Chemistry: Dye-terminator ET-amersham; 100% of reads								
	Assembly program: Phrap; version 0.990329								

Consensus quality: 140857 bases at least Q40  
Consensus quality: 144624 bases at least Q30  
Consensus quality: 146614 bases at least Q20  
Insert size: 147644; sum-of-contigs  
Quality coverage: 7.92x in Q20 bases; sum-of-

NOTE: This is a 'working draft' sequence. It currently consists of 39 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be reserved.

1	11367	11366	contig of 11366 bp in length
*	21020	20919	contig of 9553 bp in length
*	29937	29836	contig of 8817 bp in length
*	37379	37278	contig of 7442 bp in length
*	42026	41825	contig of 4547 bp in length
*	46941	46840	contig of 4815 bp in length
*	51538	51437	contig of 4497 bp in length
*	56506	56105	contig of 4961 bp in length
*	60397	60296	contig of 4091 bp in length
*	65110	65009	contig of 4613 bp in length
*	69544	69443	contig of 4334 bp in length
*	74026	73925	contig of 4382 bp in length
*	77675	77574	contig of 3549 bp in length
*	82587	82587	contig of 4913 bp in length
*	82688	82765	contig of 5178 bp in length
*	87966	87966	contig of 4105 bp in length
*	92171	96078	contig of 3908 bp in length
*	96179	100630	contig of 4452 bp in length
*	100731	103358	contig of 2628 bp in length
*	103459	107845	contig of 4387 bp in length
*	107946	110195	contig of 2250 bp in length
*	110296	111348	contig of 2853 bp in length
*	113249	116298	contig of 3050 bp in length
*	118399	120895	contig of 4997 bp in length
*	120906	123739	contig of 2744 bp in length
*	123840	126208	contig of 2369 bp in length
*	128309	128205	contig of 1897 bp in length
*	128306	132675	contig of 4370 bp in length
*	133776	134310	contig of 1535 bp in length
*	134411	137699	contig of 3289 bp in length
*	137800	139586	contig of 1787 bp in length
*	139687	141655	contig of 1969 bp in length
*	141756	143142	contig of 1387 bp in length
*	142243	144600	contig of 1358 bp in length
*	144701	146265	contig of 1565 bp in length
*	145366	147564	contig of 1199 bp in length
*	147665	148991	contig of 1327 bp in length
*	149092	150119	contig of 1028 bp in length
*	150220	151441	contig of 1222 bp in length
* NOTE: This is a 'working draft' sequence. It currently consists of 39 contigs. The true order of the pits is not known and their order in this sequence represents an arbitrary. Gaps between the contigs are represented by runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.			
*	1	11266:	contig of 11266 bp in length
*	11267	11366:	gap of 100 bp
*	11367	120919:	contig of 9553 bp in length
*	20920	21019:	gap of 100 bp
*	21020	29836:	contig of 8817 bp in length
*	29837	29936:	gap of 100 bp
*	29937	37278:	contig of 7342 bp in length
*	37279	37378:	gap of 100 bp
*	37379	41925:	contig of 4547 bp in length
*	41926	42025:	gap of 100 bp
*	42026	46840:	contig of 4815 bp in length
*	46841	46940:	gap of 100 bp
*	46941	51437:	contig of 4497 bp in length

## FEATURES

**Source**

misc\_feature

```
/note="assembly_fragment"
```



misc_feature	11367. .20919
	/note="assembly-fragment"
misc_feature	21020. .29836
	/note="assembly-fragment"
misc_feature	29937. .37278
	/note="assembly-fragment"
misc_feature	37379. .41925
	/note="assembly-fragment"
misc_feature	42026. .46840
	/note="assembly-fragment"
misc_feature	46941. .51437
	/note="assembly-fragment"
misc_feature	51538. .56105
	/note="assembly-fragment"
misc_feature	56206. .60296
	/note="assembly-fragment"
misc_feature	60397. .65009
	/note="assembly-fragment"
misc_feature	65110. .69443
	/note="assembly-fragment"
misc_feature	69544. .73925
	/note="assembly-fragment"
misc_feature	74026. .77574

Query Match	Similarity	Score	429.6	DB 2	Length	151441
Best Local	Similarity	90.4%	Pred.	No. 1e-88		
Matches	459	Conservative	0	Mismatches	49	Indels
						Gaps
						0
QY	276	acctgaactaccttgggcccccttcaacgaagcctgtactcaatccctccctgacttgggg				335
DB	26548	AGCTGAAGTACTCTGGGCCCCCTTTCACAGACCCAGACTTCAACCCCTCCCGCTGGGGG				26489
QY	335	cagaactcttgcacaggcgccacgttcaactttggaagtgtgcgaagcctcaatgaagc				395
DB	26488	CAGAACTCTGCCACAGGGCCACTGTGTGATTGGAGGTGTGGGAAGCCCAATGACAAC				26429
QY	396	tgcgcctaacccagactatgagcggtiaagtcacgtccactccctgtgttaacttgcgtgca				455
DB	26428	TTCGGCTGACCCAGAACTACGAGGCTTACAGCCACTTCTGTGTTAATTGCTGGCTCA				26369
QY	456	accgtcagagcttccacagctgtgactcgaagctagccttggcccaactctgttacagcctcc				515
DB	26368	ACCGTACAGGCTGCCACTGCTGATGCTGGCGCCAGCCTGGCCCACTTTCACACAGCTCC				26309
QY	516	agggcctctgtggcagcatttgcagttgtcavtgcagcgtcttggctaaccaatgtgccacgc				575
DB	26308	AGGGCTCTGTGGGACGACTTGGGGCGTATGAGCAGCTCTGGGCTAACCCACAGCCCAAC				26249
QY	576	ctctgcagagactagccagccttggccccctggccccctggcccaagtgactccctccaga				635
DB	26248	CGCTTCCTGTGGACTAAACCCACTTGGACTCCCTGGCCCTGCCACAGATGACTTCTCCAA				26189
QY	636	agatgatgtactcttgcctcgtcgtgaagagctgcagactcgtgtatgcgttcaagcaag				695
DB	26188	AATGTGAGCACTTCTGGCTGCTGCAAGAGACTGCACACCTGGCTGTGGCCCTGGCCAAAG				26129
QY	696	acttcaacccgcttaagaagaatgtcagcctccagcagcttcaagtcaaccttgcacttgg				755
DB	26128	ACTTCAACCGGCTCAAGAAGATGACAGCCTCCAGCACTGCAGTACACCTGCACACTGG				26069
QY	756	aggcaatgtgttctgtgactcttgacct	783			
DB	26068	GGGCTCATGGCTTGTGACTTCTTGACCTT	26041			

RESULT	10			
LOCUS	AC005849/c			
DEFINITION	AC005849	169144 bp	DNA	HTG
ACCESSION	AC005849			22-OCT-1998
VERSION	AC005849.1	GI:3779013		
KEYWORDS	HTG; HTGS_PHASE1.			

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Euryarchaeota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	Smith,D.R.
TITLE	Sequencing of Human Chromosome 10
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 169144)
AUTHORS	Smith,D.R.
TITLE	Direct Submission
JOURNAL	Submitted (22-OCT-1998) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154, USA
COMMENT	* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 506: contig of 506 bp in length
*
* 507 3033: contig of 2527 bp in length
*
* 3034 12956: contig of 9923 bp in length
*
* 12957 19521: contig of 6565 bp in length
*
* 19522 31636: contig of 12115 bp in length
*
* 31637 48670: contig of 17034 bp in length
*
* 48671 61610: contig of 12940 bp in length
*
* 61611 97329: contig of 35719 bp in length
*
* 97330 169144: contig of 71815 bp in length
*
* Location/Qualifiers

```

```

Db 143651 CGCTGGCTGGAGTGAACCACTTGACCTGCGCCCTGCCACAGTACTTCTCCAGA 143592
OY 636 agttagatgactctgtctgtctgaagagctgcagactgtctatgctgtcagcagaag 695
Db 143591 AGATGGACGACTTCTGCTGCTGAAGGAGTGCAGACTGTGCTGTGCGCCGAGG 143532
OY 696 actcaaccgcttaagaagaatgacagctccagcagcttcacacctgacttg 755
Db 143531 ACTTCAACCGCGCTCAGAAAGATGACAGCTCAGAGCTGCAGCTCAGCTGACCTGG 143472
OY 756 aggcacatggtctctgacctcgaacct 783
Db 143471 GGGCTCATGCTTCTGACTTCTGACCTT 143444

RESULT 11
AC091104/C 47958 bp DNA HTG 29-MAR-2001
LOCUS Homo sapiens chromosome 8 clone RP11-164H21 map 8, LOW-PASS
ACCESSION AC091104
SEQUENCE SAMPLING.
VERSION AC091104.1 GI:13488003
KEYWORDS HTG: HTGS-PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 47958)
TITLE Homo sapiens chromosome 8, clone RP11-164H21
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 47958)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B., Brown,A.,
Camarata,J., Campiano,A., Chang,J., Choquel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., Dearlano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., Laroque,K.,
Lamarez,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Margulis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheters,R., Meldrum,J., Menus,L.,
Mihova,T., Mlenka,V., Murphy,T., Naylor,C., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S.,
Severy,P., Sounguez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12381
Center clone name: 164_R_21

```

```

* NOTE: This record contains 57 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely

```

```

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be detected.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 777: contig of 777 bp in length
778 877: gap of 100 bp
878 1671: contig of 794 bp in length
1672 1771: gap of 100 bp
1772 2499: contig of 728 bp in length
2500 2599: gap of 100 bp
2600 3314: contig of 715 bp in length
3315 3414: gap of 100 bp
3415 4142: contig of 728 bp in length
4143 4242: gap of 100 bp
4243 5009: contig of 767 bp in length
5010 5109: gap of 100 bp
5110 5826: contig of 717 bp in length
5827 5926: gap of 100 bp
5927 6717: contig of 791 bp in length
6718 6817: gap of 100 bp
6818 7566: contig of 749 bp in length
7567 7666: gap of 100 bp
7667 8413: contig of 747 bp in length
8414 8513: gap of 100 bp
8514 9291: contig of 778 bp in length
9292 9391: gap of 100 bp
9392 10128: contig of 737 bp in length
10129 10228: gap of 100 bp
10229 10972: contig of 744 bp in length
10973 11072: gap of 100 bp
11073 11840: contig of 768 bp in length
11841 11940: gap of 100 bp
11941 12703: contig of 763 bp in length
12704 12803: gap of 100 bp
12804 13579: contig of 776 bp in length
13580 13679: gap of 100 bp
13680 14409: contig of 730 bp in length
14410 14509: gap of 100 bp
14510 15255: contig of 746 bp in length
15256 15355: gap of 100 bp
15356 16097: contig of 742 bp in length
16098 16197: gap of 100 bp
16198 16937: contig of 740 bp in length
16938 17037: gap of 100 bp
17038 17795: contig of 758 bp in length
17796 17895: gap of 100 bp
17896 18632: contig of 737 bp in length
18633 18732: gap of 100 bp
18733 19457: contig of 725 bp in length
19458 19557: gap of 100 bp
19558 20298: contig of 741 bp in length
20299 20398: gap of 100 bp
20399 21101: contig of 703 bp in length
21102 21201: gap of 100 bp
21202 21943: contig of 742 bp in length
21944 22043: gap of 100 bp
22044 22793: contig of 750 bp in length
22794 22893: gap of 100 bp
22894 23639: contig of 746 bp in length
23640 23739: gap of 100 bp
23740 24473: contig of 734 bp in length
24474 24573: gap of 100 bp
24574 25316: contig of 743 bp in length
25317 25416: gap of 100 bp
25417 26183: contig of 767 bp in length
26184 26283: gap of 100 bp
26284 27008: contig of 725 bp in length
27009 27108: gap of 100 bp
27109 27860: contig of 752 bp in length
27861 27960: gap of 100 bp

```

27961	28694:	contig of 734 bp	in length
28695	28794:	gap of 100 bp	
28795	29550:	contig of 796 bp	in length
29591	29690:	gap of 100 bp	
29691	30400:	contig of 710 bp	in length
30401	30500:	gap of 100 bp	
30501	31277:	contig of 727 bp	in length
31278	31327:	gap of 100 bp	
31328	32051:	contig of 724 bp	in length
32052	32151:	gap of 100 bp	
32152	32874:	contig of 723 bp	in length
32875	32974:	gap of 100 bp	
32975	33714:	contig of 740 bp	in length
33715	33814:	gap of 100 bp	
33815	34586:	contig of 772 bp	in length
34587	34686:	gap of 100 bp	
34687	35444:	contig of 758 bp	in length
35445	35544:	gap of 100 bp	
35545	36288:	contig of 724 bp	in length
36289	36368:	gap of 100 bp	
36369	37118:	contig of 750 bp	in length
37119	37218:	gap of 100 bp	
37219	37948:	contig of 730 bp	in length
37949	38048:	gap of 100 bp	
38049	38709:	contig of 741 bp	in length
38790	38889:	gap of 100 bp	
38890	39618:	contig of 729 bp	in length
39619	39718:	gap of 100 bp	
39719	40469:	contig of 751 bp	in length
40470	40569:	gap of 100 bp	
40570	41306:	contig of 737 bp	in length
41307	41406:	gap of 100 bp	
41407	42164:	contig of 758 bp	in length
42165	42264:	gap of 100 bp	
42265	43003:	contig of 739 bp	in length
43004	43103:	gap of 100 bp	
43104	43837:	contig of 734 bp	in length
43838	43937:	gap of 100 bp	
43938	44555:	contig of 618 bp	in length
44556	44655:	gap of 100 bp	
44656	45391:	contig of 736 bp	in length
45392	45491:	gap of 100 bp	
45492	46344:	contig of 743 bp	in length
46345	46344:	gap of 100 bp	
46335	47105:	contig of 771 bp	in length
47106	47205:	gap of 100 bp	
47206	47958:	contig of 753 bp	in length

FEATURES	SOURCE
Location/Qualifiers	1. .47958
/organism="Homo sapiens"	
/dp_xref="taxon:9606"	
/chromosome="8"	
/map="8"	
/clone="RP11-164H21"	
/clone_lib="RPCT-11	Human Male BAC
11146 a 9513 c 8428 g 12707 t 6164 others	

Query Match	5.6%	Score 46.2	DB 2	Length 47958
Best Local Similarity	49.6%	Pred. No. 0.62		
Matches 117	Conservative	0	Mismatches 119	Indels
			Gaps	0
QY 15	cgcgcgagccgcgcgtgcctcccatccacgtctgtgagagagccgcgcgg	74		
Db 33607	GCCCGCGCCGCGCGCCGCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCGG	33548		
QY 75	cgcgcgcgcgcgcgcgcgcacatgaccctgcgagcagcgagatcgtggggagatgaatt	134		
Db 33547	CCCCGCGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC	33488		
QY 135	gcctatgacggtgtgtgtgcacatcccttcgacgtccttaatgcacaagagatc	194		

[illegible]

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----
Center project name: H.NH0440608
----- Summary Statistics -----
Sequencing vector: M13; 81%
Sequencing vector: Plasmid; 19%
Chemistry: Dye-terminator Big Dye; 78% of reads
Chemistry: Dye-terminator Big Dye; 22% of reads
Assembly program: Phrap; version 0.990119
Consensus quality: 173977 bases at least Q40
Consensus quality: 177014 bases at least Q30
Consensus quality: 179042 bases at least Q20
Insert size: 179000; agarose-1p
Insert size: 184627; sum-of-contigs
Quality coverage: 5.08 in Q20 bases; agarose-1p
Quality coverage: 5.01 in Q20 bases; sum-of-contigs

NOTE: This is a working draft sequence. It currently
* consists of 32 contigs. The true order of the contigs
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1126: contig of 1126 bp in length
1127
1127: gap of unknown length
1227
1227: contig of 2172 bp in length
3398
3398: gap of unknown length
3399
3399: gap of unknown length
5591
5591: contig of 2093 bp in length
5592
5592: gap of unknown length
7340
7340: contig of 1648 bp in length
7440
7440: gap of unknown length
9562
9562: contig of 2123 bp in length
9663
9663: gap of unknown length
12135
12135: gap of unknown length
12235
14073: contig of 1839 bp in length
14074
14074: gap of unknown length

```

*	14114	1669:	contig of 2496 bp in length
*	16670	1660:	gap of unknown length
*	16770	1860:	contig of 1831 bp in length
*	18601	18700:	gap of unknown length
*	18701	22418:	contig of 3718 bp in length
*	22419	22518:	gap of unknown length
*	22519	26559:	contig of 4041 bp in length
*	26560	26659:	gap of unknown length
*	30860	30895:	contig of 436 bp in length
*	30966	30995:	gap of unknown length
*	34750	34749:	contig of 3754 bp in length
*	34849	34849:	gap of unknown length
*	34850	39440:	contig of 4591 bp in length
*	39441	39540:	gap of unknown length
*	39541	45183:	contig of 3643 bp in length
*	43184	43283:	gap of unknown length
*	43284	47528:	contig of 4245 bp in length
*	47529	47628:	gap of unknown length
*	47629	51967:	contig of 4339 bp in length
*	51968	52067:	gap of unknown length
*	52068	58827:	contig of 6760 bp in length
*	58828	58927:	gap of unknown length
*	58928	63567:	contig of 4640 bp in length
*	63568	63667:	gap of unknown length
*	63668	68120:	contig of 5453 bp in length
*	69121	69220:	gap of unknown length
*	69221	73215:	contig of 3995 bp in length
*	73216	73316:	gap of unknown length
*	73316	79585:	contig of 6270 bp in length
*	79586	79685:	gap of unknown length
*	79686	89651:	contig of 9966 bp in length
*	89652	89751:	gap of unknown length
*	89752	95856:	contig of 6105 bp in length
*	95857	95956:	gap of unknown length
*	95957	102877:	contig of 6921 bp in length
*	102878	102977:	gap of unknown length
*	110335	110335:	contig of 7358 bp in length
*	110336	110435:	gap of unknown length
*	110436	110472:	contig of 10037 bp in length
*	120473	120572:	gap of unknown length
*	120573	132066:	contig of 11494 bp in length
*	132067	132166:	gap of unknown length
*	132167	144622:	contig of 12456 bp in length
*	144623	144722:	gap of unknown length
*	144723	156389:	contig of 11667 bp in length
*	156390	156489:	gap of unknown length
*	156490	169023:	contig of 12534 bp in length
*	169024	169123:	gap of unknown length
*	169124	187727:	contig of 18604 bp in length

	misc_feature	/note="assembly_name:Contig29"	18701..22418
	misc_feature	/note="assembly_name:Contig30"	22519..26559
	misc_feature	/note="assembly_name:Contig31"	26660..30895
	misc_feature	/note="assembly_name:Contig32"	30996..34749
	misc_feature	/note="assembly_name:Contig33"	34850..39440
	misc_feature	/note="assembly_name:Contig34"	39541..43183
	misc_feature	/note="assembly_name:Contig35"	43284..47528
	misc_feature	/note="assembly_name:Contig36"	47629..51967
	misc_feature	/note="assembly_name:Contig37"	52068..58827
	misc_feature	/note="assembly_name:Contig38"	58928..63567
	misc_feature	/note="assembly_name:Contig39"	63668..69120
	misc_feature	/note="assembly_name:Contig40"	69221..73215
	misc_feature	/note="assembly_name:Contig41"	73316..79585
	misc_feature	/note="assembly_name:Contig42"	79686..89651
	misc_feature	/note="assembly_name:Contig43"	89752..95856
	misc_feature	/note="assembly_name:Contig44"	95957..102877
	misc_feature	/note="assembly_name:Contig45"	102978..110335
	misc_feature	/note="assembly_name:Contig46"	110436..120472
	misc_feature	/note="assembly_name:Contig47"	clone_end:r7
	misc_feature	/note="assembly_name:Contig48"	vector_side:right"
	misc_feature	/note="assembly_name:Contig49"	132167..144622
	misc_feature	/note="assembly_name:Contig50"	144723..156389
	misc_feature	/note="assembly_name:Contig51"	156490..169023
	misc_feature	/note="assembly_name:Contig52"	clone_end:sp6
	misc_feature	/note="assembly_name:Contig53"	vector_side:right"
	misc_feature	/note="assembly_name:Contig54"	169124..187727
BASE COUNT	43736 a	48685 c	47674 g 44349 t 3283 other
ORIGIN			
Query Match	5.4%	Score 44.4	DB 2; Length 187727;
Best Local Similarity	51.3%	Pred. No. 1.1;	
Matches 156; Conservative	0;	Mismatches 141; Indels	7; Gaps 2;
OY	276	acctgaactcttgggcccctttcaagagcgcgtacatcatcttccttccttgaggggg	335
Db 103565	AACGCTCAGTACAGGGCAACCCCTTAGTGACCCTGGCGTTCTTCAGGCCCTGTAGACTTCACAC	103624	
OY	336	cagaactctgcccgaggccacgcagctcaactcttgaaagtgttgcgaagaccataaacagagc	395
Db 103625	TCAAGAGCGCTGCTCTCTGCACACGGCTTTTAAGAAGCTGGCACGGCCGTGAATGACGGGG	103684	
OY	396	tgcgcgtcacccagacaattgagc-----gtacagtcacctctctgttaactttgcgtgac	451
Db 103685	AACGGCTGAGCTTACCCAGAGGGCATTTGACACCAACGACACTCTCAGCTCTGTATAGACGAC	103744	
OY	452	c---tcaacgctcagcgcttcacacagctgacacgcagacgttagcctggccccacttcgttac	508
Db 103745	CAGAGCGACTTAACCTTGACAGTCCATCTCTGCGGGCTACAGCTCGGGGCTCGAGACACT	103804	



/chromosome="16"  
/map="16p13.3"  
/tissue-type="lung"  
1.1357  
/gene="NHERF-2"  
1.1014  
/gene="NHERF-2"  
/note="PDZ-like protein; closely related to NHERF; also known as E3KARP and TKA-1; potential splice variant of SIP-1"  
/codon\_start=1  
/product="Na+/H+ exchanger regulatory factor 2"  
/protein\_id="AAC52090.1"  
/db\_xref="GI:2665826"  
/translation="MAPEPLRPLRLVGEQGYGFHLHGKGRGQFTRVPPGSP  
AAEALRAGRLVENVNGVNEGTTHQVORIKAVEGOTRLVVDDELRRLRRLT  
CTEEMARGLPRAHPDPEPKPDMAHTSHSSEAGKRVSPRLRLRHLKRGPG  
YGNLHSDKSRPGQYIRSVDPGSPAASGLRADRLLEVNGVNGEGLRAHVAASIKR  
REDEARLLVDPETDEHFKRLRVPTVEHVGGLPSVVTGTSPOALNGSACSRSR  
LPSSDKTDEGSAMKODPFQESGLHLSPTAAEKERKARAVKRAPOMDMNRREIF  
SNF"

BASE COUNT 267 a 440 c 470 g 180 t  
ORIGIN

Query Match 5.2%; Score 42.4; DB 9; Length 1357;  
Best Local Similarity 52.9%; Pred. No. 11;  
Matches 91; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 512 ctccagggcctgctgggagcagatgcatgctgctgagagcttgctaccacatgcc 571  
|||||  
Db 286 ctccggccggcgagcagctgctgacgagagatggcccgagagcctgctccacccgcc 345  
|||||  
QY 572 cagcctctgccaaggaactgagcagcctgagccctgagccctgccaagtgactctc 631  
|||||  
Db 346 CAGACCCCTGGAGAGCCGAGACTGGGACACACACGCGACACAGCTCCGAAGCT 405  
|||||  
QY 632 cagaagatgagatgctctgctgctgagagagctgagacactgctatgac 683  
|||||  
Db 406 GGCAGAGAGATGTCACTGGGCCCTGAGGAGAGCTGCGCCCTGCTCTGCC 457  
|||||

RESULT 15  
AF004900 1524 bp mRNA PRI 07-OCT-1998  
LOCUS Homo sapiens NHE3 kinase A regulatory protein E3KARP mRNA, complete cds.  
ACCESSION AF004900  
VERSION AF004900.1 GI:2198848  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1524)  
Yun,C.H., Oh,S., Zizak,M., Steplock,D., Tsao,S., Tse,C.M.,  
Weilman,E.J. and Donowitz,M.  
cAMP-mediated inhibition of the epithelial brush border Na+/H+  
exchanger, NHE3, requires an associated regulatory protein  
Proc. Natl. Acad. Sci. U.S.A. 94 (7), 3010-3015 (1997)  
97250481  
JOURNAL MEDLINE  
REFERENCE 2 (bases 1 to 1524)  
Yun,C.H., Lampecht,G., Forster,D.V. and Sidor,A.  
NHE3 kinase A regulatory protein E3KARP binds the epithelial brush  
border Na+/H+ exchanger NHE3 and the cytoskeletal protein ezrin  
J. Biol. Chem. 273 (40), 25856-25863 (1998)  
98421507  
JOURNAL MEDLINE  
REFERENCE 3 (bases 1 to 1524)  
Yun,C.H.  
Direct Submission  
TITLE Submitted (20-MAY-1997) Medicine, Johns Hopkins School of Medicine,  
GI Unit, 918 Ross Building, 720 Rutland Avenue, Baltimore, MD  
21205, USA

FEATURES  
source Location/Qualifiers  
1.1524  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
17.1030  
/note="NHE3 associated regulatory protein, NHE3 kinase A  
regulatory protein"  
/codon\_start=1  
/product="E3KARP"  
/protein\_id="AAC63061.1"  
/db\_xref="GI:2198849"  
/translation="MAPEPLRPLRLVGEQGYGFHLHGKGRGQFTRVPPGSP  
AAEALRAGRLVENVNGVNEGTTHQVORIKAVEGOTRLVVDDELRRLRRLT  
CTEEMARGLPRAHPDPEPKPDMAHTSHSSEAGKRVSPRLRLRHLKRGPG  
YGNLHSDKSRPGQYIRSVDPGSPAASGLRADRLLEVNGVNGEGLRAHVAASIKR  
REDEARLLVDPETDEHFKRLRVPTVEHVGGLPSVVTGTSPOALNGSACSRSR  
LPSSDKTDEGSAMKODPFQESGLHLSPTAAEKERKARAVKRAPOMDMNRREIF  
SNF"

BASE COUNT 305 a 498 c 495 g 226 t  
ORIGIN

Query Match 5.2%; Score 42.4; DB 9; Length 1357;  
Best Local Similarity 52.9%; Pred. No. 11;  
Matches 91; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 512 ctccagggcctgctgggagcagatgcatgctgctgagagcttgctaccacatgcc 571  
|||||  
Db 302 ctccggccggcgagcagctgctgacgagagatggcccgagagcctgctccacccgcc 361  
|||||  
QY 572 cagcctctgccaaggaactgagcagcctgagccctgagccctgccaagtgactctc 631  
|||||  
Db 362 CAGACCCCTGGAGAGCCGAGACTGGGACACACACGCGACACAGCTCCGAAGCT 421  
|||||  
QY 632 cagaagatgagatgctctgctgctgagagagctgagacactgctatgac 683  
|||||  
Db 422 GGCAGAGAGATGTCACTGGGCCCTGAGGAGAGCTGCGCCCTGCTCTGCC 473  
|||||

Search completed: March 18, 2002, 10:26:51  
Job time: 3202 sec

Mon Mar 18 11:23:40 2002

us-09-931-704-4.rge





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2002, 09:32:34 ; Search time 1206.01 Seconds

(Without alignments)  
7297.446 Million cell updates/sec

Title: US-09-931-704-4

Perfect score: 819

Sequence: 1 tatattaaagcttcgcgcg.....agccacagtcagctgtgctt 819

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:\*

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	544	66.4	1053	11	BG164929
2	543	66.3	853	11	BG437538
3	476	58.1	476	10	BE632644
4	464.4	56.7	955	10	AL543945
5	333.6	40.7	512	10	AI390475
6	312.4	38.1	407	10	AI752561
7	261.2	31.9	488	10	AA015243
8	246.8	30.1	552	11	BG095271
9	228.4	27.9	440	11	BG148676
10	172.8	21.1	915	11	BF035982
11	124.8	15.2	691	11	BF213570
12	66.4	8.1	180	10	AA204015

13	50.8	6.2	711	11	BG758365
14	50.2	6.1	925	13	CNS0091P
15	45	5.5	925	13	CNS0091P
16	43.4	5.3	936	13	CNS01SR2
17	43.4	5.3	1101	13	CNS01258
18	42.8	5.2	844	13	CNS0052P
19	42.8	5.2	932	13	CNS0072Q
20	42.4	5.2	505	10	AI984845
21	42.4	5.2	658	10	AW873554
22	42.4	5.2	690	11	BF732971
23	42.4	5.2	787	11	BF971932
24	42.4	5.2	983	11	BF337781
25	42	5.1	839	13	CNS004NB
26	41.6	5.1	511	10	BE754133
27	41	5.0	342	11	BI345660
28	41	5.0	778	13	AC327192
29	41	5.0	1203	13	CNS015V4
30	40.8	5.0	525	11	BE808147
31	40.8	5.0	603	10	AI039288
32	40.8	5.0	1171	13	BI0823
33	40.6	5.0	844	13	CNS0052P
34	40.4	4.9	268	10	AW355114
35	40.4	4.9	776	13	CNS010RY
36	40.2	4.9	797	13	CNS04NGE
37	40.2	4.9	828	13	CNS04GHZ
38	40	4.9	935	13	CNS006XK
39	39.6	4.8	586	10	AV670279
40	39.4	4.8	467	10	AW657443
41	39.4	4.8	529	10	BE033291
42	39.4	4.8	978	10	BE614877
43	39.4	4.8	1009	13	CNS010EW
44	39	4.8	1281	11	BG852363
45	38.8	4.7	995	10	AL538217

## ALIGNMENTS

RESULT 1

LOCUS BG164929 1053 bp mRNA EST

DEFINITION 60234355F1 NIH\_MGC\_89 Homo sapiens cDNA clone IMAGE:4453813 5', mRNA sequence.

ACCESSION BG164929

VERSION BG164929.1 GI:12671563

KEYWORDS EST.

SOURCE Human.

ORGANISM Homo sapiens

REFERENCE NIH-MGC <http://mgc.nhl.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution info: <http://imgc.llnl.gov>  
Found through the I.M.A.G.E. Consortium/LLNL at:  
Plate: L1AM10244 row: m column: 14  
High quality sequence start: 3  
High quality sequence stop: 675.  
Location/Qualifiers  
1..1053  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4453813"  
/clone\_lib="NIH\_MGC\_89"  
/tissue\_type="Hypermephroma, cell line"







LOCUS	AI752861	407 bp	mRNA	EST	22-JUN-1999
DEFINITION	cn17d05.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA				
ACCESSION	AI752861				
VERSION	AI752561.1 GI:5130825				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 407) Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Hotchkiss, R.N. and Franccomano, C.A. SCAP: The Skeletal Anatomy Project Unpublished (1997)				
TITLE	Contact: Libin Jia				
JOURNAL	Medical Genetics Branch				
COMMENT	National Human Genome Research Institute 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA Tel: 301-402-4877 Fax: 301-496-7157 Email: libin@helix.nih.gov DNA Sequencing and analyses by National Institutes of Health Intramural Sequencing Center (NISC). Plate: 17 row: d column: 05 Seq primer: -21M13 forward primer (ABI).				

FEATURES	Location/Qualifiers
SOURCE	1. .407

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NH7BC-cn17d05"
/clone_lib="Normal Human Trabecular Bone Cells"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI"

```

Query Match	38.1%	Score 312.4	DB 10	Length 407
Best Local Similarity	90.3%	Pred. No. 8.1e-59		
Matches 334	Conservative 0	Mismatches 36	Indels 0	Gaps 0
Oy	414	atgaagcgtaacgttaactctctgttttacttgcgtgtgcctaaacgtaagctgcacag	473	
Db	9	ACGAGGCGCTACAGCCACTTCGTGTACTTCGCGGCTCCACCGCTCAGCGTGCACAG	68	
Oy	474	ctgaactccgaagtatgacctgcggcccaactctgtaccagacctccaggccctgcgtggcagca	533	
Db	69	CTGACTCTGCGCGCAGCTTGCGCCCACTTCTGCACCAAGCCTCCAGGCGCTGCTGGGCAGCA	128	
Oy	534	ttgcagcgtgtcatalgcyagcgtttgcttaaccacactgcccaagcctctgcacgaagctagc	593	
Db	129	TTGCGGCGGTATGCGACACTCTGGGCTACCCACTGCCGCCAGCGCGCTCGTGAGATGAAC	188	
Oy	594	cagccttggccccctgcgcctcccaacagttaattccatccagaagatgatactctgcgc	653	
Db	189	CCACTTTGGAGCTCTTGCCCTGCCACAGAGATTCCTCCAGAAAGATGGACGACTTCTGGAC	248	
Oy	654	tgcctaagaagcgtgcagacctgtgcatatgagcgttccagccaaagacttcaacgcgcttaaga	713	
Db	249	TTCTTAAGAGACTGCGAGACTTGCTGTGTGGCGCTCGGCCCAAGACTTCACCCGGCTCAGA	308	
Oy	714	agaagaatgcagcctccagcaagcttcaatgaacctgcacttggaggcacatagtttctgcac	773	
Db	309	AGAAAGTGCAGCGCTCCAGCAGCTGCAGTACACCTGCAGACTGGGGGCTATGCTTCTGAC	368	
Oy	774	ctctgcacct 783		
Db	369	TTCTGACCTT 378		

RESULT	7
AA015243	
LOCUS	488 bp mRNA EST J.JAN-1997
DEFINITION	mh3oc11.r1 Soares mouse placenta 4NbMp13.5 14.5 M; muscleus cDNA
KEYWORDS	clone IMAGE:444020 5', mRNA sequence.
ACCESSION	AA015243
VERSION	AA015243.1 GI:1476292
SOURCE	EST.
ORGANISM	house mouse.

ORGANISM	Mus musculus
Eukaryota:	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia:	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 488)
AUTHORS	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubugue, T., Geisell, S., Kuwaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellengberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE	The WashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through IMI; contact the  
IMI Consortium (info@image.llnl.gov) for further information.  
MGI:269356  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 479.  
Location/Qualifiers  
1. 488

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/cdoe="IMAGE:44020"
/clone_lib="Soares mouse placenta 4NtMP] 14.5"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pT73(1) (Pharmacia)
with a modified polylinker; Site_1: Not ; Site_2: Eco RI
1st strand cDNA was primed with a Not I - oligo(dT) primer
15'
TCCTCCCAATCTGAGGTGGAGCGGCCCGCGGAATTCTTTTCTTTTTTTTT
T T 3'; double-stranded cDNA was ligated to Eco RI adaptor
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M.Patino Honaldo."

```

[illegible]

	Db	175	GGGATGTAGTCTGGCTTAAGCAGCGTGCTGTGGCA-CTCCGTCAGACTGCCAGTTCTTAA	233
Oy	182	cgcaacaggaagatccaggccttgccccctccatccagaataactatgacctaacccgtacc	241	
Db	234	CGCACAGGAAGATCCAGGCCCTGGCCCCCTCCATCCAGAAAACTATGACCTCACCCGCTAC	293	
Oy	242	ctgagacatcaactccgcagcttagctggaaactaact	279	
Db	294	CTGGAGCATCAACTCCGCAGCTTACTGGGACTTACGT	331	
RESULT	8	BG095271	mRNA	EST
LOCUS	BG095271/c	u85q07.x1 Soares mouse NMBG.bcell Mus musculus cDNA clone		26-JAN-2001
DEFINITION	IMAGE:3383604.3.	similar to TF:Q9QZM3 Q9QZM3 NEUROTROPHIN-1/B-CELL STIMULATING FACTOR-3., mRNA sequence.		
ACCESSION	BG095271			
VERSION	BG095271.1	GI:12577834		
KEYWORDS	EST.			
SOURCE ORGANISM	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 552) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov This clone is available royalty-free through LNLN ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:1088600			
JOURNAL COMMENT				
FEATURES				
SOURCE				
		location/qualifiers		
		1..552		
		/organism="Mus musculus"		
		/db_xref="taxon:10090"		
		/clone="IMAGE:3383604"		
		/clone_1id="Soares_mouse.NMBG.bcell"		
		/lab_host="DH10B (phage-resistant)"		
		/note="Organ: germinal B-cell; Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCACATCTGAAGTGGGAGCGCGCCGCTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library is normalized; constructed by Bento Soares and M.Patrina Bonaldo."		
BASE COUNT	147 a 125 c 147 g 131 t	2 others		
ORIGIN				
Query Match	30.1%;	Score 246.8;	DB 11;	Length 552;
Best Local Similarity	96.3%;	Pred. No. 2.1e-44;		
Matches 262;	Conservative 0;	Mismatches 9;	Indels 1;	Gaps 1.
Oy	548	ggcagcgttggctacccaactgccaccagtccctgcacaagactgaacagcctggcccct	607	
Db	552	GCGATGCTGGCTACCCACTGCCCTCTTTGGCAGAGGACTGACCCAGCCCTGGGCCCT	493	
Oy	608	ggccctggcccaagtgacttcctccagaagatgtagcttctgctctgtaagaagctg	667	
Db	492	GGCCCTGCCACAGTGAATTCTCCAGAAGATGAGACTTCTGGCTG-TTAAAGAGCTG	434	
Oy	668	cagagccggcatgcygtctcagccaagactccaacggcgttaaagaagaagattgacgct	727	
Db	433	CAGACCTGGCTATGGCGTTACAGCAAGACACTTCAACCGCGCTTAATAAAGAATGACGCT	374	
Oy	728	ccagcagcttcaagtaeccctcgaacttggaggcacatglttcttgaccttgaaccttaac	787	

Db	373	CCAGACGCTTGAGTACACCTGCACCTTGGAGGCCCATGATTGACACTTTGACCTTAAC	314
Qy	788	ccccacacctccagggccagtcagctgtgctt	819
Db	313	ccccacacctccagggccagtcagctgtgctt	282
RESULT	9		
LOCUS	BGI148676	440 bp	mRNA
DEFINITION	u85907.y1 Soares_mouse_MNGB.bcel1 Mus musculus cDNA clone IMAGE:3383604 5' similar to T8:Q9QZM3 Q9QZM3 NEUROPROPHIN-1/B-CELL STIMULATING FACTOR-3. ;, mRNA sequence.		
ACCESSION	BGI148676		
VERSION	BGI48676.1	GI:12652098	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	NCI-CCAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Other ESTs: u85907.x1 Contact: Robert Strausberg, Ph.D. Email: cgapb-remail.nih.gov This clone is available royalty-free through LNL. Contact the IMAGE Consortium ( <a href="mailto:InfoImage.LNL.gov">InfoImage.LNL.gov</a> ) for further information. Seq primer: -40RP from Glibco MGI:1088600		
FEATURES	High quality sequence stop: 431.		
SOURCE	Location/Qualifiers		
	1..440		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:3383604"		
	/clone_id="Soares_mouse_MNGB_bcel1"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: germinal B-cell; Vector: 1: (3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGGTCCCAATCTGAAGTGGAGGCGGCCCGCTGTTTTTTTTTTTTTTTTTTT 3'] ; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized; constructed by Benito Soares and M.Petina Bonaudo."		
BASE COUNT	99 a 127 c 108 g 106 t		
ORIGIN			
Query Match	27.9%; Score 228.4; DB 11; Length 440;		
Best Local Similarity	99.6%; Pred. No.2.3e-40;		
Matches	229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	590	gagccagcttgaggcccttgagcccttgccacagctgacttccctcagaagaatgagacttc	649
Db	1	GAGCCAGCCTGGGCCCCCTGGCCCTGCCACAGATGACTTCCTCCAGAGATGATGACTTC	60
Qy	650	tgagctctgaaaggaagctcagagctcagctggtatagggattagccaaagacttcaaccgactt	709
Db	61	TGGCTGCTGAAAGAGACTTCAGACTGGGCTATAGGGCTTCAAGCCAAAGACTTCAACGGGCTT	120
Qy	710	aagaagaagatgacagctccacagcagcttcagctacccttcgacacttgagagcaatgattc	769
Db	121	AAGAAGAAGATGACAGCCTTCAGACACTTCAGTACACCTTGAGAGGCCCATGCTTC	180
Qy	770	tgacctctgaccttaaccccccaacctccagagccagctcagctgtgctt	819













PR 30-JAN-1998; 9805-0016534.  
 PR 03-FEB-1997; 9705-0792019.  
 XX (AMGE-) AMGEN INC.  
 PA Chang M, Elliot GS, Sarmiento U, Senaldi G;  
 PI WPI: 1998-437475/37.  
 DR P-PSDB; AAM29716.  
 XX  
 XX Newly isolated nucleic acid encoding human or murine neurotrophic  
 PT factor NNT-1 - useful for treatment of neurological and  
 PT immunological diseases or inflammation, also as vaccine adjuvant  
 PS  
 PS Disclosure: Fig 4; 120p; English.  
 XX  
 XX This newly isolated mouse cDNA sequence codes for a novel  
 CC neurotrophic factor, designated NNT-1 (see AAM29716), that is a  
 CC growth factor for neurons and for B or T cells. Vectors and  
 CC host cells are provided for use in the production of murine and  
 CC human NNT-1 polypeptides. These are used to treat: (i) neurological  
 CC or immunological diseases, specifically Alzheimer's, Parkinson's or  
 CC Huntington's diseases, amyotrophic lateral sclerosis,  
 CC Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy and  
 CC degeneration of the neural retina, or conditions characterised by T  
 CC or B cell defects, e.g. common variable immunodeficiency (CVID),  
 CC selective IgA deficiency, hypogammaglobulinaemia and X-linked  
 CC agammaglobulinaemia (claimed), but many others disclosed; and (ii)  
 CC inflammation. NNT-1 is also able to boost immunoreactivity and  
 CC antibody production following vaccination, and, since it inhibits  
 CC tumour necrosis factor production, it may also be useful for  
 CC treating sepsis. NNT-1 nucleic acid fragments are also used as  
 CC hybridisation probes in diagnostic assays. In addition, cells that  
 CC have been engineered to express NNT-1 can be implanted, or nucleic  
 CC acids are delivered in gene therapy vectors.  
 XX  
 XX Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 819; DB 19; Length 819;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-196;  
 Matches 819; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 421 gtaacagtaacctctgtgtacttgcgtgagctcaaccgctcaagctctgcaact 480  
 QY 481 ccgagctaaccttggccactctgtaccagccccaaggcctctggcagacttgag 540  
 DB 481 ccgagctaaccttggccactctgtaccagccccaaggcctctggcagacttgag 540  
 QY 541 tctcagcagactctgtctaccacttgcctccagcctctgccaagagcttgagcact 600  
 DB 541 tctcagcagactctgtctaccacttgcctccagcctctgccaagagcttgagcact 600  
 QY 601 ggcctctggcctctgcccacagctgactctccccaagaatgagatgactctgtct 660  
 DB 601 ggcctctggcctctgcccacagctgactctccccaagaatgagatgactctgtct 660  
 QY 661 ggcctctggcctctgcccacagctgactctccccaagaatgagatgactctgtct 720  
 DB 661 ggcctctggcctctgcccacagctgactctccccaagaatgagatgactctgtct 720  
 QY 721 gcaagctcagcagcttcaatcactcctgcaacttggaggaatgtttcttactct 780  
 DB 721 gcaagctcagcagcttcaatcactcctgcaacttggaggaatgtttcttactct 780  
 QY 781 ccttaaccccaacactccagagcccaagctcaagctgtgctt 819  
 DB 781 ccttaaccccaacactccagagcccaagctcaagctgtgctt 819  
 DE  
 RESULT 2  
 AAV22654  
 ID AAV22654 standard; cDNA; 819 BP.  
 XX  
 XX AAV22654;  
 AC  
 XX 13-JUL-1998 (first entry)  
 DT  
 XX  
 XX cDNA encoding murine neurotrophic factor NNT-1.  
 DE  
 XX  
 XX Mouse: neurotrophic factor; NNT-1; growth; motor; synu; ctic; neuron;  
 KW treatment; neurological disease; degeneration; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; ALS; Alzheimers's disease; stroke; ss.  
 XX  
 XX Mus sp.  
 OS  
 XX  
 XX Key 95..771 Location/Qualifiers  
 FH CDS  
 FT /\*tag= a 95..175  
 FT sig\_peptide /\*tag= b 176..769  
 FT mat\_peptide /\*tag= c  
 FT  
 XX  
 XX US5741772-A.  
 PN  
 XX 21-APR-1998.  
 PD  
 XX  
 XX 03-FEB-1997; 9705-0792019.  
 PF  
 XX 03-FEB-1997; 9705-0792019.  
 PR  
 XX (AMGE-) AMGEN INC.  
 PA  
 XX Chang M;  
 PI  
 XX WPI: 1998-260526/23.  
 DR P-PSDB; AAM56142.  
 XX  
 XX Neurotrophic factor NNT-1 polypeptide and related nucleic acids -  
 PT useful for stimulating growth of motor and sympathetic neurons  
 PS  
 PS Disclosure: Fig 4; 41pp; English.  
 CC The present sequence encodes a murine neurotrophic factor, designated  
 CC NNT-1, which is capable of stimulating growth of motor or sympathetic

CC neurons. The NNT-1 protein is useful in the treatment of neurological diseases characterised by the degeneration and death of particular CC classes of neurons. These diseases specifically include Parkinson's disease, amyotrophic lateral sclerosis (ALS), Alzheimer's disease, stroke and various degenerative disorders affecting vision.

SQ Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 other;

Query Match 100.0%; Score 819; DB 19; Length 819;  
Best Local Similarity 100.0%; Pred. No. 2.2e-196;  
Matches 819; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tattatataagcttcgagcgagcgagcgcctccacccacccagcagcctcggaga 60  
DB 1 tattatataagcttcgagcgagcgagcgcctccacccacccagcagcctcggaga 60  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 120  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 120  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 180  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 180  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 240  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 240  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 300  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 300  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 360  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 360  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 420  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 420  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 480  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 480  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 540  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 540  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 600  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 600  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 660  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 660  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 720  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 720  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 780  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 780  
QY gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 819  
DB gagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 819

RESULT 3  
AAA39483  
ID AAA39483 standard; cDNA; 819 BP.

XX  
AC AAA39483;  
XX  
DT 24-AUG-2000 (first entry)  
XX  
DE Murine NNT-1 cDNA.  
XX  
KW NNT-1; neurotrophic factor; neuroprotective; treatment;  
KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;  
KW nervous system degeneration; Alzheimer's disease; Parkinson's disease;  
KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome; murine;  
KW Huntington's disease; peripheral neuropathy; neural retina degeneration;  
KW retinopathy; immune disorder; hematopoietic disorder; ss.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 95..772  
FT /\*tag= a  
FT /product= "NNT-1"  
XX  
PN US6054294-A.  
XX  
PD 25-APR-2000.  
XX  
PF 12-DEC-1997; 97US-0988819.  
XX  
PR 03-FEB-1997; 97US-0792019.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Chang M;  
XX  
DR WP1; 2000-338492/29.  
DR P-PSDB: AAY87814.  
XX  
PT New nucleic acids encoding neurotrophic factors useful for stimulating  
PT growth of motor or sympathetic neurons for treating neuron cell damage  
XX  
PS Claim 2a; Fig 4; 42pp; English.  
XX  
CC This invention describes a novel nucleic acid molecule (I) encoding a  
CC novel neurotrophic factor (NNT-1) (II) which has neurotrophic,  
CC neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and  
CC optalmimological activity. (I) is useful for producing NNT-1  
CC polypeptides which are useful for treating patients in whom various  
CC cells of the central, autonomic, or peripheral nervous system have  
CC degenerated and/or have been damaged by congenital disease, trauma,  
CC mechanical damage, surgery, stroke, ischemia, infection, metabolic  
CC disease, nutritional deficiency, malignancy and/or toxic agents. NNT-1  
CC proteins are used to treat diseases like Alzheimer's, Parkinson's,  
CC amyotrophic lateral sclerosis, Charcot-Marie-Tooth syndrome, Huntington's  
CC disease, peripheral neuropathy induced by diabetes or other metabolic  
CC disorders, and/or dystrophies or degeneration of the neural retina such  
CC as retinitis pigmentosa, drug-induced retinopathies, stationary forms of  
CC night blindness; progressive cone-rod degeneration, immune disorders and  
CC hematopoietic disorders. (I) is effective in treating neurological  
CC conditions and promotes neuron regeneration. Neural functions are  
CC effectively restored in patients suffering from various neurological  
CC disorders. This sequence encodes the murine NNT-1 protein described in  
CC the method of the invention.  
XX  
SQ Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 other;

Query Match 100.0%; Score 819; DB 21; Length 819;  
Best Local Similarity 100.0%; Pred. No. 2.2e-196;  
Matches 819; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tattatataagcttcgagcgagcgagcgcctccacccacccagcagcctcggaga 60  
DB 1 tattatataagcttcgagcgagcgagcgcctccacccacccagcagcctcggaga 60

















Db	300	caggggccactcttctgactcttggaaggtctgtgacgaagcctcaatgacaactctgcgtgaccca	359
QY	409	gaactatgaaggcgtacagcttcacctcctctgttlaacttgcgtggccctcaaccgttaagctgc	468
Db	360	gaactacgaaggcctcaacagccacccctctctgttgaactctcgtgctcaaccgtcaagctgc	419
QY	469	cacagctggaactccgaacgtagcctctgcccacactctgtaccaagcctcaaggctcgtctgg	528
Db	420	cactgtctgaagctgcgcgcgacgctgcgcacactcttgacacagcctccaaggctcgtctgg	479
QY	529	cagcatgtgcaggtgtctaattgagaaagcttgtgtaacccactgcccacagctcttgccaaggac	588
Db	480	cagcctctgcggcgctcaatgtgaagctctcgtgggtacccactgcaccaagcctctgctggagac	539
QY	589	tgaagcaagcctggggccctctggccctctgcacaaagatctctctccagaagaatgatgacct	648
Db	540	tgaaccacactctggaactcctctggcccttgcccaagaatgaaactctctccagaagaatgaaagact	599
QY	649	cttgcgtctgaagaagagctctgcagacactctgtctatgtgcgttccagccaaggaacttcaacgcgt	708
Db	600	cttgcgtctgaagaagagctctgcagacactctgtctgtgcgtctggccaaggaacttcaacgcgt	659
QY	709	taagaagaagaatgcagcctccagcagcttcaagttcaaccttgtaacttggaagcaaatggtt	768
Db	660	caagaagaagaatgcagcctccagcagcttcaagttcaaccttgtaacttggaagcaaatggtt	719
QY	769	ctgacacttgaacct 783	
Db	720	ctgactcttgacct 734	

RESULT	12
AAD04201	ID
AAD04201	standard; cdna; 729 BP.
XX	XX
XX	AC
XX	AAD04201;
XX	XX
DT	02-JUL-2001 (first entry)
XX	XX
DE	Human cardiotrophin-like cytokine (CLC) cdna.
XX	XX
KW	Human; biologically active complex; haemopoietin receptor; NR6;
KW	cardiotrophin-like cytokine; CLC; therapy; prophyllaxis; proliferation;
KW	differentiation; cell survival; neurotrophic activity; ss.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	CDS
FT	11..679
FT	/**tag- a
FT	/product= "Human cardiotrophin-like cytokine (CLC)
FT	protein"
FT	/note="CDS does not include stop codon"
FT	/partial
FT	11..91
FT	/**tag- b
FT	92..679
FT	/**tag- c
FT	/product= "Human mature cardiotrophin-like cytokine
FT	(CLC) protein"
XX	XX
PN	WO200127157-A1.
XX	XX
PD	19-APR-2001.
XX	XX
PF	06-OCY-2000; 2000MO-AU01216.
XX	XX
PR	08-OCY-1999; 99AU-0003327.
PR	12-MAY-2000; 2000AU-0007489.
XX	XX
PA	(AMRA-) AMRAD OPERATIONS PTY LTD.
XX	XX

PI Nash A., Uachino KM, Fabri LJ, Reid K, Bartlett PF, H. Leon DJ,  
 PI Nakata Y., Hasegawa M,  
 XX WPI: 2001-281978/29.  
 DR P-PSDB; AAE00828.  
 XX  
 PT New biologically active complex comprising NR6 and  
 PT cardiotrophin-1-like-cytokine, for facilitating proliferation,  
 PT differentiation and/or survival of a cell -  
 XX  
 PS Claim 31; Page 112-114; 123pp; English.  
 XX  
 CC The present invention relates to a biologically active complex comprising  
 CC a haemopoietin receptor, NR6 and cardiotrophin-1-like cytokine (CLC).  
 CC The complex is useful in the manufacture of a medicament for the  
 CC treatment and/or prophylaxis of a subject, as it is involved in  
 CC facilitating proliferation, differentiation and/or survival of a cell.  
 CC The complex or its components have neurotrophic activity. The present  
 CC sequence is human cardiotrophin-1like cytokine (CLC) cDNA.  
 XX  
 SO Sequence 729 BP; 132 A; 261 C; 196 G; 140 T; 0 other;

Query Match	73.3%	Score	600.2	DB	22	1	729
Best Local Similarity	91.6%	Pred. No.	1.8e-141				
Matches 635; Conservative	0	Mismatches	58	Indel	2	0	Gaps
						0	

OY	91	ccccatgacctccgaacagagggactcgttggggatgtttgacttgcctatcttcaggatct	150
Db	7	ccccatgacctccgaacagagggactcgttggggatgtttgacttgcctatcttcaggatct	66
OY	151	gtggcaacctccctcagttgcagagctcttaatctgaacagggatccaaagcccttggccctc	210
Db	67	cttggaacctccctcagttgcagagctctcaatctgaacagggagcccaagggtcttggccctc	126
OY	211	catccagaanaacctatgacctcaaccgcgtacctgtgagatcaatccctccgaacttaacgt	270
Db	127	catccagaanaacctatgacctcaaccgcgtacctgtgagatcaatccctccgaacttaacgt	186
OY	271	gaactactgaactgaactctggggcccccttccaacagagacgtgacttcaatccctcctgact	330
Db	187	gaactactgaactgaactctggggcccccttccaacagagacgtgacttcaatccctcctgact	246
OY	331	gggggcagaanaacctgcccgaagggccacagttcaactcttgagttgttgagaaatcaatga	390
Db	247	gggggcagaanaacctgcccgaagggccacactctgtgactcttgagttgttgagaaatcaatga	306
OY	391	caaggctcgggtctgaaccagaactatgaagcggtatgaatgacatccctctgtttttcttgcgtg	450
Db	307	caaatctgggtctgaaccagaactatgaagcggtatgaatgacatccctctgtttttcttgcgtg	366
OY	451	cctcaaacccgttaagctctgcacaaagtctgaacctccgaagctgaagctctggcccaactcttgataccag	510
Db	367	cctcaaacccgttaagctctgcacaaagtctgaagctctggccgaagctctggcccaactcttgataccag	426
OY	511	cctccaagggcctctgtcgggcagagctatgaaattgatactgacagctcttgagcttccatctggcc	570
Db	427	cctccaagggcctctgtcgggcagagctatgaaattgatactgacagctctcttgagctcttccatctggcc	486
OY	571	ccagagctctctcgaaggagatggaaccagagctctggccccctggccctctgaccccaagctgaacttccct	630
Db	487	ccagagctctctcgaaggagatggaaccacacttgaaactcctctggccctctgaccccaagctgaacttccct	546
OY	631	ccagaagaatgtgaatgaactcttctgctgcctggaaggagctctgacaaactggtatcttcttcagcag	690
Db	547	ccagaagaatgtgaatgaactcttctgctgcctggaaggagctctgacaaactggtatcttcttcagcag	606
OY	691	caaggaactcaaccggctcttaagaagaagatggaagctcccgagcctcagacgtctcaatctcaacccttga	750
Db	607	caaggaactcaaccggctctcaagaagaagatggaagctcccgagcctcagacgtctcagatctcaacccttga	666
OY	751	ctctggaagcacaatgattctgaaccttgacctctgaacctctctgaacctctctgaacctctctgaacct	810



```

FT      exon      /tag= 9
FT      exon      /number=2
FT      exon      /tag= h
FT      exon      /number= 3
PN      WO9839922-A1.
PD      06-AUG-1998.
PF      02-FEB-1998: 98WO-US02363.
PR      03-JAN-1998: 98US-0016534.
PR      03-FEB-1997: 97US-0792019.
XX      (AMGE-) AMGEN INC.
XX      Chang M, Elliot GS, Sarmiento U, Senaldi G;
XX      WPI: 1998-437475/37.
XX      P-PSDB: AAM29715.
XX      Newly isolated nucleic acid encoding human or murine neurotrophic
XX      factor NNT-1 - useful for treatment of neurological and
XX      immunological diseases or inflammation, also as vaccine adjuvant
XX      Claim 4; Fig 2; 120pp; English.
XX      This newly isolated human genomic DNA sequence (deposited at
XX      ATCC 98294) codes for a novel neurotrophic factor, designated NNT-1
XX      (see AAM29715), that is a growth factor for neurons and for B or T
XX      cells. It was obtained from a human genomic PL library using the
XX      human NNT-1 cDNA (see AAV47510) as probe. Vectors containing the
XX      cDNA or genomic DNA and host cells are provided for use in the
XX      production of NNT-1 polypeptides. These are used to treat: (1)
XX      neurological or immunological diseases, specifically Alzheimer's,
XX      Parkinson's or Huntington's diseases, amyotrophic lateral sclerosis,
XX      Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy and
XX      degeneration of the neural retina, or conditions characterized by T
XX      or B cell defects, e.g. common variable immunodeficiency (CVID),
XX      selective IGA deficiency, hypogammaglobulinemia and X-linked
XX      agammaglobulinemia (claimed), but many others disclosed; and (11)
XX      inflammation. NNT-1 is also able to boost immunoreactivity and
XX      antibody production following vaccination, and since it inhibits
XX      tumour necrosis factor production, it may also be useful for
XX      treating sepsis. NNT-1 nucleic acid fragments are also used as
XX      hybridisation probes in diagnostic assays. In addition, cells that
XX      have been engineered to express NNT-1 can be implanted, or nucleic
XX      acids are delivered in gene therapy vectors.
XX      Sequence 5088 BP; 992 A; 1746 C; 1191 G; 1158 T; 1 other:
SQ

```

```

Query Match      52.5%; Score 429.6; DB 19; Length 5088;
Best Local Similarity 90.4%; Pred. No. 1.9e-98;
Matches 459; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```

```

QY      276 accgtgaactcctggggccccccttcaacgagcgtactcaatccctccgagctggg 335
      1 |||||
DB      3364 agctgaactactctgggccccttcaacgagcgtactcaatccctccgagctggg 3423
QY      336 cagaagaactcgcagcagcgtcaacttggaagtgtggcgaagcctcaatgaagc 395
      |||||
DB      3424 cagagactcgcagcagcgtcaacttggaagtgtggcgaagcctcaatgaagc 3483
QY      396 tgcggtgacccagcaactatgagcgtagcagtcacccctgtgttacttgcggtgctca 455
      |||||
DB      3484 tgcggtgacccagcaactatgagcgtagcagtcacccctgtgttacttgcggtgctca 3543
QY      456 accgttcagctgcagcagcgtcaactcgaagcgtgagcgtgagcgtgagcgtgctcc 515
      |||||
DB      3544 accgttcagctgcagcagcgtcaactcgaagcgtgagcgtgagcgtgagcgtgctcc 3603

```

```

QY      516 aggcgcctgctggcagcagcatttgagtgatcagcagcgttgagcagcagcagc 575
      |||||
DB      3604 aggcgcctgctggcagcagcatttgagtgatcagcagcgttgagcagcagcagc 3663
QY      576 ccttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 635
      |||||
DB      3664 cgcctgcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3723
QY      636 agatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 695
      |||||
DB      3724 agatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3783
QY      696 actcaacccgcttaagaagaatcagcagcagcagcagcagcagcagcagcagcagcagc 755
      |||||
DB      3784 actcaacccgcttaagaagaatcagcagcagcagcagcagcagcagcagcagcagcagc 3843
QY      756 aggcacatggttctgacccctgaacct 783
DB      3844 gggcctatggtcttctgacccctgaacct 3871

```

## RESULT 15

```

AAV22653
ID      AAV22653 standard; DNA: 5088 BP.
XX
AC      AAV22653;

```

```

DT      13-JUL-1998 (first entry)

```

```

DE      Human genomic DNA encoding neurotrophic factor NNT-1.

```

```

XX      Human: neurotrophic factor; NNT-1; growth; motor; sympathetic neuron;
XX      treatment; neurological disease; degeneration; Parkinson's disease;
XX      amyotrophic lateral sclerosis; ALS; Alzheimer's disease; stroke; ss.
XX      Homo sapiens.

```

```

OS      Homo sapiens.

```

```

FH      Key

```

```

FT      misc-feature

```

```

FT      Location/Qualifiers

```

```

FT      138

```

```

FT      /*tag= a

```

```

FT      /note= "represents intervening unsequenced region

```

```

FT      of 1 Kb"

```

```

FT      US5741772-A.

```

```

FT      21-Apr-1998.

```

```

FT      03-FEB-1997; 97US-0792019.

```

```

FT      03-FEB-1997; 97US-0792019.

```

```

FT      (AMGE-) AMGEN INC.

```

```

FT      Chang M;

```

```

FT      WPI: 1998-260526/23.

```

```

FT      Neurotrophic factor NNT-1 polypeptide and related nucleic acids

```

```

FT      useful for stimulating growth of motor and sympathetic neurons

```

```

FT      Disclosure; Fig 2; 41pp; English.

```

```

SQ

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2002, 09:37:40 ; Search time 44.41 Seconds  
(without alignments)  
4176.660 Million cell updates/sec

Title: US-09-931-704-4  
Perfect score: 819  
Sequence: 1 tatataaagcttcgccg.....agcccgatcgctgctt 819

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCBUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	819	100.0	819	1	US-08-792-019B-4
2	819	100.0	819	3	US-08-988-819-4
3	819	100.0	819	3	US-09-016-534-4
4	669.4	81.7	797	1	US-08-792-019B-1
5	669.4	81.7	797	3	US-08-988-819-1
6	669.4	81.7	797	3	US-09-016-534-1
7	627	76.6	1710	3	US-09-106-182-1
8	429.6	52.5	5087	1	US-08-792-019B-3
9	429.6	52.5	5087	3	US-08-988-819-3
10	429.6	52.5	5087	3	US-09-016-534-3
11	52	6.3	396	3	US-09-106-182-7
12	43.2	5.3	6803	3	US-08-665-259-19
13	43.2	5.3	6803	3	US-08-762-500-19
14	42.4	5.2	1642	2	US-08-665-037-1
15	42.4	5.2	1642	2	US-08-666-067-1
16	42.4	5.2	1642	2	US-08-732-870-1
17	38.4	4.7	68750	3	US-09-335-409-1
18	34.8	4.2	4403765	4	US-09-103-840A-2
19	34.8	4.2	4411529	4	US-09-103-840A-1
20	34.4	4.2	1225	1	US-08-197-496A-1
21	34.4	4.2	1225	1	US-08-275-370A-1
22	33.6	4.1	13875	2	US-08-734-344-1
23	33.6	4.1	36519	3	US-08-923-137-2
24	33.4	4.1	1117	4	US-09-347-819-5
25	33.2	4.1	1600	3	US-08-602-791-1
26	33.2	4.1	3386	4	US-09-041-886-1
27	33	4.0	297	4	US-08-905-223-115

C 28	33	4.0	1226	1	US-08-367-968-1	Sequence 1, Appl
C 29	33	4.0	1226	1	US-08-665-484-1	Sequence 1, Appl
C 30	32.6	4.0	4897	6	5196516-7	Patent No. 5196516
C 31	32.4	4.0	548	4	US-09-286-132-3	Sequence 3, Appl
C 32	32.4	4.0	1397	3	US-08-188-930-231	Sequence 21, App
C 33	32.4	4.0	35060	3	US-08-814-095-7	Sequence 7, Appl
C 34	32.2	3.9	1664	3	US-09-339-993-1	Sequence 1, Appl
C 35	32.2	3.9	4080	2	US-08-446-345-35	Sequence 16, Appl
C 36	32	3.9	856	3	US-09-188-930-16	Sequence 4, Appl
C 37	32	3.9	5337	2	US-08-784-512-4	Sequence 4, Appl
C 38	32	3.9	5337	2	US-09-176-228-4	Sequence 2, Appl
C 39	32	3.9	71989	4	US-09-443-501A-2	Sequence 2, Appl
C 40	31.8	3.9	2881	1	US-08-396-479B-5	Sequence 3, Appl
C 41	31.8	3.9	2881	1	US-08-818-823-5	Sequence 1, Appl
C 42	31.6	3.9	3226	3	US-08-870-126-10	Sequence 10, Appl
C 43	31.6	3.9	745	1	US-08-036-555B-163	Sequence 163, App
C 44	31.6	3.9	745	1	US-08-469-569-163	Sequence 163, App
C 45	31.6	3.9	745	1	US-08-249-322A-163	Sequence 163, App

## ALIGNMENTS

RESULT 1  
US-08-792-019B-4  
; Sequence 4, Application US/08792019B  
; Patent No. 5741772  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, MING-SHI  
; TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: AMGEN INC.  
; STREET: 1840 DEHAVILLAND DRIVE  
; CITY: THOUSAND OAKS  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 91320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/792,019B  
; FILING DATE: 03-FEB-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOK, ROBERT R.  
; REGISTRATION NUMBER: 31,602  
; REFERENCE/DOCKET NUMBER: A-442  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 819 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 95..769  
; FEATURE:  
; NAME/KEY: mat.peptide  
; LOCATION: 176..769  
; FEATURE:  
; NAME/KEY: sig.peptide  
; LOCATION: 95..175  
; NAME/KEY: LOCATON:  
; LOCATION: 95..175  
; US-08-792-019B-4

Query Match 100.0%; Score 819; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 6e+206;  
Matches 819; Conservative 0; Mismatches 0; Gaps 0;

QY 1 tatattaagcttcgacgagccgagctcgccctccacccacccgagcttgagaga 60  
DB 1 TATATTAAGCTTCGCGGAGCGCGGCTCCGCTCCGCTCCGCTTGGAGA 60  
QY 61 gggagccgcccggccgcccggcccccagcccaatgagcttcgagcagggagctg 120  
DB 61 GGAAGCCG 120  
QY 121 gggagatgtagcttcgacatgacggtgctgtgacacccctcgagcttcagctta 180  
DB 121 GGGGATGTTAGCTTGCCTATGACAGGCTGCTGTGGACCTCCGCGAGTCTTAA 180  
QY 181 tggcagaagagatccagggccctggccctccatccagaaacctatgacctaccgcta 240  
DB 181 TCGCAGAGGAGATCCAGGCGCTGGCGCTCCGCGCGCGCGCGCGCGCGCG 240  
QY 241 ccggagacatcaatccgagcttagctggagcttcagcttcagcttcgagcttc 300  
DB 241 CCGGAGCATCAATCCGAGCTTACCTGAGCTAGCTGAACTGAGCTGAGCTTT 300  
QY 301 caagagctgagcttcacatccctccctcctcctcctcctcctcctcctcctc 360  
DB 301 CAACGAGCTGAGCTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
QY 361 caactggagatgtagcttcgagagcttcagagcttcgagcttcgagcttcgag 420  
DB 361 CAACTTGGAGATGAGCTGAGAGCTTCAATGAGCTGAGCTGAGCTGAGCT 420  
QY 421 gtagagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 480  
DB 421 GTACAGTCACT 480  
QY 481 cgaagctgagcttcgagacatctctcctcctcctcctcctcctcctcctcct 540  
DB 481 CGAGCTGAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
QY 541 tgcatagctcagagcttcgagcttcagcttcagcttcagcttcagcttcagct 600  
DB 541 TGTATGCGGAGAGCTTGGCTTACCTCCGCTCCGCTCCGCTCCGCTCCGCT 600  
QY 601 gggccctgagcttcgagacatctcctcctcctcctcctcctcctcctcctc 660  
DB 601 GGGCCCTGAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
QY 661 ggaagctcagagcttcgagcttcagcttcagcttcagcttcagcttcagct 720  
DB 661 GGAAGCTCAGAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
QY 721 ggaagctcagagcttcagcttcagcttcagcttcagcttcagcttcagct 780  
DB 721 GGAAGCTCAGAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
QY 781 ccttaagcttcagagcttcagcttcagcttcagcttcagcttcagcttcag 840  
DB 781 CCTTAAGCTTCAGAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840

## RESULT 2

US-08-988-819-4  
Sequence 4, Application US/08988819  
Patent No. 6054294  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: NEUTROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,819  
FILING DATE: 12-DEC-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442A  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 819 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 95..769  
FEATURE:  
NAME/KEY: mat-peptide  
LOCATION: 176..769  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: 95..175  
US-08-988-819-4  
Query Match 100.0%; Score 819; DB 3; Length 819;  
Best Local Similarity 100.0%; Pred. No. 6e-206;  
Matches 819; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 tatattaagcttcgacgagccgagctcgccctccacccacccgagcttgagaga 60  
DB 1 TATATTAAGCTTCGCGGAGCGCGGCTCCGCTCCGCTCCGCTTGGAGA 60  
QY 61 gggagccgcccggccgcccggcccccagcccaatgagcttcgagcagggagctg 120  
DB 61 GGAAGCCG 120  
QY 121 gggagatgtagcttcgacatgacggtgctgtgacacccctcgagcttcagctta 180  
DB 121 GGGGATGTTAGCTTGCCTATGACAGGCTGCTGTGGACCTCCGCGAGTCTTAA 180  
QY 181 tggcagaagagatccagggccctggccctccatccagaaacctatgacctaccgcta 240  
DB 181 TCGCAGAGGAGATCCAGGCGCTGGCGCTCCGCGCGCGCGCGCGCGCGCG 240  
QY 241 ccggagacatcaatccgagcttagctggagcttcagcttcagcttcgagcttc 300  
DB 241 CCGGAGCATCAATCCGAGCTTACCTGAGCTAGCTGAACTGAGCTGAGCTTT 300  
QY 301 caagagctgagcttcacatccctccctcctcctcctcctcctcctcctcctc 360  
DB 301 CAACGAGCTGAGCTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
QY 361 caactggagatgtagcttcgagagcttcagagcttcgagcttcgagcttcgag 420  
DB 361 CAACTTGGAGATGAGCTGAGAGCTTCAATGAGCTGAGCTGAGCTGAGCT 420  
QY 421 gtagagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 480  
DB 421 GTACAGTCACT 480  
QY 481 cgaagctgagcttcgagacatctcctcctcctcctcctcctcctcctcctc 540  
DB 481 CGAGCTGAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540



ADDRESSEE: AMGEN INC.  
STREET: 1840 DEHAVILLAND DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/792,019B  
FILING DATE: 03-FEB-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 797 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 90..764  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 171..764  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 90..170  
US-08-792-019B-1

Query Match 81.7% Score 669.4; DB 1; Length 797;  
Best Local Similarity 92.0%; Fred. No. 9.9e-167;  
Matches 717; Conservative 0; Mismatches 61; Indels 1; Gaps 1;  
QY 5 attaagctcgccgagccgagctgcgcctccaccatccgcagcctctctggagagag 64  
DB 1 ATTAAGCTTGCGCGAGCCCGGGCTCCGCCCTCCACTCCGACGCTCCGGAGAGAG 60  
QY 65 ccgagcccgccgagcccgagcccccagcccatggaacctccgagcagggagactg 124  
DB 61 CCGACCCCGCGCGCCAG-CCCCAGCCCATGACCTCCGAGCAGGGGACTCGTGGGG 119  
QY 125 atgttaagctcatgagcagctgtgtgagacctccctctgacagctcttaatcgc 184  
DB 120 ATGTTAGGCTCTGTGACAGGTGCTGTGGACCTCTCTGAGTCCAGTCTCAATGCG 179  
QY 185 acagagagatccagagccctgagccctccatccagaaacctatgacctaccgcct 244  
DB 180 ACAGGAGACCCAGGGCCCTGGCCCTCATTCAGAAACCTATGACTACCCGCTACTG 239  
QY 245 gagatacaactccgagcttgcctgaggaactaccgaactctggggccctttcaac 304  
DB 240 GAGCACCACACTCCGACCTTGCTGGGACCTATCTGAACCTACCTGGGCCCTTTCAAC 299  
QY 305 gaagctgactcaatctctctgactgagggcagaaactctgcccagggcagagtcac 364  
DB 300 GAGCAGACCTTCAACCTCCCGCGCTGGGGCAGACACTCTGCCAGGGCCACTGTGAC 359  
QY 365 ttggaagctgtggcgaagcctcaatgacagcgtgcgctgagcccaactatgagcgctac 424  
DB 360 TTGGAAGCTGTGGCAGACCTCAATGACAACTGGCGCTGAGCCCACTAGAGAGCTTAC 419  
QY 425 agtcacactctgtgtaactgtgtgtgacctaacacgtcagctgacagctgaactcga 484  
DB 420 AGCCACCTTTGTGTTACTGTGCGTGCCTCAACCGTACGAGCTGCCTGACTGCGC 479

QY 485 cgtagcctgcccacactctctgtaccagcctccagggcctgctgagcagratccaggtg 544  
DB 480 CGAGCCCTGGCCCACTTCTGTGACACGAGCTCCAGGGCTCTGTGGAGAGATTGCCGGGCTC 539  
QY 545 atggcagcctgtgcttaccacactgcccagcctctgtccaggaagtcagaccacgtgg 604  
DB 540 ATGGCACTCTGGGCTACCCACTGCCAGCCGCTGCTGGAGCTGAACCCACTTTGACT 599  
QY 605 cctgcccctgcccagctgacttctctcagaagaatgagtaactctgtctgtcgaagag 664  
DB 600 CCTGGCCCTGCCCACTGACTTCTCCAGAGATGAGACACTTGTGCTGTGAAGAG 659  
QY 665 ctgcagacctgtgctatgagcgttccagccaaggaactcaacgggcttaagaaqjgatcag 724  
DB 660 CTGCAGACCTGTGGCTGTGGCCTCGCCCAAGACTTCAACCGGCTCAAGAGAAGATGACG 719  
QY 725 cctccagcagcttcaatcaccctgacttggaggaatggttctgaccttgacct 783  
DB 720 CCTCCAGCACTGACATCACCCTGACCTGGGGCTCATGTGACTTGTGACTTGCACCTT 778

RESULT 5  
US-08-988-819-1  
Sequence 1, Application US/08988819  
Patent No. 6054294  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,819  
FILING DATE: 12-DEC-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442A  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 797 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 90..764  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 171..764  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 90..170  
US-08-988-819-1

Query Match 81.7% Score 669.4; DB 3; Length 797;



```

Db      480  CGAGCCTGGCCACTCTTGACACGACGCTCCAGGCGCTGCTGGGACACATTCGGGGCCTC 539
QY      545  atggcaagcttggtctacccaatgccccagcctctgccaggactgagccagcctgggccc 604
Db      540  ATGGCAAGCTTGGGCTACCCACCTGCCCCAGCGCGCTGGCTGAGTGAACCCACTTGGACT 599
QY      605  ccggcgcctcccaaatgactctcctcagaagaatgactctgctgctgctgaagag 664
Db      600  CCGGGCCCTCCACACTGACTTCTCCAGAGTGAACACTTCTGCTGCTGCTGAAGAG 659
QY      665  ctgcaagcttggtatggttcaagcaagactcaacggcttaagaaagaatgag 724
Db      660  CTCAGACCTGGGCTGTGGCGCTGGCCAGAGACTTCAACCGGCTCAGAGAAGATGACAG 719
QY      725  cctcagcaagcttcaagcaccctgagccttgaggagacatgcttctgactctgacct 783
Db      720  CCTCAGACGACTCAGTACCTGACCTGACCTGGGGCTCATGCTTGTGACTTGCACCTT 778

```

## RESULT 7

US-09-106-182-1

Sequence 1, Application US/09106182

Patent No. 6046035

GENERAL INFORMATION:

APPLICANT: Shi, Yanggu

APPLICANT: Ruben, Steve

TITLE OF INVENTION: Cardiotrophin-Like Cytokine

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc

STREET: 9410 Key West Ave

CITY: Rockville

STATE: MD

COUNTRY: US

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/106,182

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/051,053

FILING DATE: 30-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF985

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1710 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 46..720

FEATURE:

NAME/KEY: sig-peptide

LOCATION: 46..126

FEATURE:

NAME/KEY: mat-peptide

LOCATION: 127..720

US-09-106-182-1

Query Match 76.6%; Score 627; DB 3; Length 1710;  
 Best Local Similarity 91.7%; Pred. No. 1,66-155;  
 Matches 674; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

```

QY      49  gctctggagagagagagagagagagagagagagagagagagagagagagagagagagag 108
Db      1  GCTTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 59
QY      109  agggagactcgttgggagatgttagcttgcatacagagtgctgtgtgtgtgtgtgtgtgt 168
Db      60  AGGGAGACTCGTGGGGAGATGTAGAGTGTGCTGTGACAGGTGCTGTGCACTTCCTGCAGT 119
QY      169  gccagctctaattgcacagagagatccagagcctgtgccccctcatalcagaanaactatga 228
Db      120  GCCAGCTCTAATGCGACAGAGGAGCCAGAGGCGCTGGGCCCTCCATCCAGAAACCTATGA 179
QY      229  cctcaccgccttaacttgagatcaactccagagcttagcttggagacactgactgactact 288
Db      180  CTTACCCGCTACTCTGAGACCACTCCGACGCTTGGCTGGGACTATCTGAACCTACT 239
QY      289  ggggcccccttcaagagccttgactcaatcctcctcagactggggagagatcctgccc 348
Db      240  GGGCCCCCTTTCAACAGAGCAGACCTTCAACCCCTCCCGCTGGGGGACAGAACTGTGCC 299
QY      349  cagggcacagtgtaacttggaagtgtgagagagcctaatagacagctgcggtgtgacca 408
Db      300  CAGGGCCACTGTGACTTGGAGGTGGCGAAGGCTTCMAATGACMAATGCCCGCTGAGACCA 359
QY      409  gaactatgagcgctacagtaactcctcctgttacttgcgtgagcctcaacagctgagctgc 468
Db      360  GAACCTAGAGAGGCTACAGCCACCTTCTGTGTACTTGTGCTGAGCTCAACCTTAGCTGCT 419
QY      469  cacagctgaactcgcagctagcttgagccactctctgacagcctccagagatgtgtgtgg 528
Db      420  CACTGCTGAGCTGCGCGCGCTGGGCCACTTCTGACACAGCCTCCAGAGATTCGTGGG 479
QY      529  cagcatgacagtgatcagtgagcagccttgctacccactcccaagcctctgtgtgtgtgt 588
Db      480  CACGATTTGGCGGGGTGATGAGCAGCTGTGGGCTGACCTCCCACTCCCAAGCTGCTGGGAC 539
QY      589  tgaagcagcctgggcccccttgagcccccaagtgactctcctccagaagaatgagtaact 648
Db      540  TGAAGCCACTTGGACTCTGGCCCTGCCCTCCACAGTACTTCTCCAGAAAGATGACACT 599
QY      649  ctgctgctgaagagagagagagagagagagagagagagagagagagagagagagagagag 708
Db      600  CTGGCTGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
QY      709  taagaagaagatgacagctccagagagagagagagagagagagagagagagagagagagag 768
Db      660  CAAGAAGAAGATGACAGCTCCAGACAGCTGAGTACAGCTGAGTACAGCTGAGTACAGCTGAG 719
QY      769  ctgagctctgacct 783
Db      720  CTGACTCTGACCTT 734

```

## RESULT 8

US-08-792-019B-3

Sequence 3, Application US/08792019B

Patent No. 5741772

GENERAL INFORMATION:

APPLICANT: CHANG, MING-SHI

TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: AMGEN INC.

STREET: 1840 DEHAVILLAND DRIVE

CITY: THOUSAND OAKS

STATE: CA

COUNTRY: USA

ZIP: 91320

Query Match	52.5%	Score 429.6	DB 1	Length 5087
Best Local Similarity	90.4%	Pred. No. 1.3e-103		
Matches 459; Conservative	0;	Mismatches 49;	Indels 0;	Gaps 0

RESULT 9  
US-08-988-819-3  
; Sequence 3, Application US/089888815  
; Patent No. 6054294

Query Match	52.5%;	Score 429.6;	DB 3;	Length 5087;
Best Local Similarity	90.4%;	Pred. No. 1,3e+10;		
Matches 459;	Conservative	0;	Mismatches 49;	Indel. 0;
				Gaps 0

QY	276	accgtaactaactygggcccctttaaagaagccgtaactaaipccctccatctgagcgaag	335
Db	3363	AGCTGAACTACTCTGGGGCCCCCTTTCAAGAGCCAGACTTCAACCTCTCCCGCTGGGGG	3422
QY	336	cagaaactctggccaggcgcaagctcaacttggaaagtgtggaaagcctcaatgaagcgc	395
Db	3423	CAGAACTCTTCCAGGGCCACTGTGACTTGGAGGTGGGGAAGCTCAATGACAAC	3482
QY	396	tgcgactgaaccagaactatgaagcgttaagtaactcctctgtgtacttccatggccctca	455
Db	3483	TCCGCTGACCCAGAACTACAGAGGCCATCAGCCACCTTGTGTTACTTGGCTGGCTCA	3542
QY	456	accgtcaagctgcgcacagcttgaactccgaagtatgcttgagccactctcgtatctgaagccctcc	515
Db	3543	ACCGTCAGAGCTGCACACTGCTGACTGTGGCGCCGACCTGGGCCCACTTTCGACCAAGCTTCACC	3602
QY	516	aggagcctctgagcagcatltagcgtgtcaatggcagaagcttgctaccacactgcccagc	575
Db	3603	AGGGCTGCTGGGGCAGCATTTCCGGGGCTCATGCGACAGCTGGGGCTACCCGATTTCCAGACG	3662
QY	576	ctctgcagaggaactgaagccagcctctgggcccctggccctggcccacaaatgaattccctccaga	635
Db	3663	CGCTGCTGGGACTTAAACCACTTGGACTCCTGGCCCTGCCACAACTGACTTTCTTCAGA	3722
QY	636	agatgtagtaactctgagctgtctgaagaagagctgcagaaactctgctatgtgcatltagcgtltagccaaag	695
Db	3723	AGATTGACACACTTCTGGCTGCTGAAGAAGAGCTGACGACTGCTGGCTGGCGCTCTCCCAAGG	3782

QY 696 actcaaccggttaagaagaagatgcagctccagcagcttaagtcacccctgactgg 755  
|||||  
Db 3783 ACTTCAACCGGCTCAAGAAACATGAGCTCCAGCAGCTGACGATCACCTGCG 3842  
QY 756 aggcacatggtctctgacctgacct 783  
|||||  
Db 3843 GGGCTCATGGCTTCTGACTTGTGACCTT 3870

## RESULT 10

US-09-016-534-3  
; Sequence 3, Application US/09016534  
; Patent No. 6153874  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, MING-SHI  
; APPLICANT: ELLIOTT, GARY S.  
; APPLICANT: SARMIENTO, ULIA  
; APPLICANT: SENALDI, GIORGIO  
; TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC.  
; STREET: ONE AMGEN CENTER  
; CITY: THOUSAND OAKS  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 91320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,534  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/792,019  
; FILING DATE: 03-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOK, ROBERT R.  
; REGISTRATION NUMBER: 31,602  
; REFERENCE/DOCKET NUMBER: A-442B  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5087 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 137..138  
; OTHER INFORMATION: /product= "INTERVENING UNSEQUENCED"  
; OTHER INFORMATION: REGION OF >1KB"  
US-09-016-534-3

Query Match 52.5%; Score 429.6; DB 3; Length 5087;  
Best Local Similarity 90.4%; Pred. No. 1.3e-103;

Matches 459; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 276 acctgaactcctggggcccccttcaacgagcctgactcaatcctcctcgaactgggg 335  
Db 3363 AGCTGAACCTACTGGGCCCCCTTCAACGAGCCAGACTTCAACCCCTCCCGCTGGGGG 3422  
QY 336 cagaacctctcccgagggcccgctgcaacttggaagtgtggcgaagcctcaatgacagc 395  
Db 3423 CAGAGACTCTGCGCCAGGGCCACTGTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482  
QY 396 tggcggtgaccacgaactatgagcgctacagtcacccctctgttacttggtgctctca 455  
|||||

Db 3483 TCGGGCTGACCCAGAACTACGAGGCTTACAGCCACTTCTGTGTACTTCA 3542  
QY 456 accgtcaggtctgcacagctgaactcagcagctgtagcctgaccactctgtaacagcctcc 515  
Db 3543 ACCGTGAGGCTGCGCCACTCTGAGCTGCGCCGCGAGCTGCGCCACTTCTGTGTACTTCA 3602  
QY 516 aggcctgctgagcagctgcaatgcatgagtgaggaagcctggtctacccactgcccagc 575  
Db 3603 AGGGCTGCTGGGCAAGCATTTGCGGGCTGATGAGCTGCTGGGCTTACCTGCGCCAGC 3662  
QY 576 ctctgcaggggactgagcagcctgggccccctggccccctgcccacagtgatctccacaga 635  
Db 3663 CGGTGCTGGGAGCTGAGAACCCACTTGTGACTCTGGCCCTGCGCCAGCTTCTGTGAGA 3722  
QY 636 agatgatgactctgtgctgctgtagagagctgagcagctgcatgagctgcaagaag 695  
Db 3723 AGATGAGAGACTTCTGCTGCTGTAAGAGAGCTGAGAGCTGAGACTGCTGCTGCGCCAGG 3782  
QY 696 actcaaccggtctaaagaagaatgagcctcagcctcagcagctcagtcacccctgactgg 755  
Db 3783 ACTTCAACCGGCTCAAGAAACATGAGCTCCAGCAGCTGACGATCACCTGCG 3842  
QY 756 aggcacatggtctctgacctgacct 783  
Db 3843 GGGCTCATGGCTTCTGACTTGTGACCTT 3870  
|||||

## RESULT 11

US-09-106-182-7  
; Sequence 7, Application US/09106182  
; Patent No. 6046035  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Yangu  
; APPLICANT: Ruden, Steve  
; TITLE OF INVENTION: Cardiotrophin-Like Cytokine  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc  
; STREET: 9410 Key West Ave  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/106,182  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/051,053  
; FILING DATE: 30-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PF385  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 396 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-106-182-7

Query Match 6.3%; Score 52; DB 3; Length 390;





Db 6588 CCCCCCTTCTCTAGCCTGCGGCGGCGCTTCTCTGCGACGCGCGCTGCGGCCCG 6529

QY 204 gccctccatccagaacattagaccctaccgctactgag 247

Db 6528 CCCCCCGGAGCTGTCGCTGTCGCGCCGACCGCGGAGAGCGCGCG 6485

RESULT 14  
US-08-665-037-1

; Sequence 1, Application US/08665037

; Patent No. 5895813

; GENERAL INFORMATION:

; APPLICANT: Seedorf, Klaus

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT

; TITLE OF INVENTION: OF TKA-1 RELATED

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/665,037

; FILING DATE: June 13, 1996

; CLASSIFICATION: 530

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 60/005,167

; FILING DATE: October 13, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 220/156

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1642 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 97...1446

; US-08-665-037-1

Query Match 5.2%; Score 42.4; DB 2; Length 1642;  
Best Local Similarity 52.9%; Pred. No. 0.03; Mismatches 81; Indels 0; Gaps 0;  
Matches 91; Conservative 0;

QY 512 ctccagagcctctgagcagcatgagtgatgagcagcttgctaccactgccc 571

Db 379 CTCGCCCGCGGCGAGCTGTACCGAGAGATGGCCAGGAGGGCTCCACCGCGC 438

QY 572 cagcctctccagagactgagccagcctgagccctgagccctccacagtgactctc 631

Db 439 CAGGACCCCTGGAGCGCAAGCAGACTGGGCGACACACCGGCGACCAACAGCTCGAAGCT 498

QY 632 cagaagatgatgactctgctgctgaggaagactgagacactgctgctatgac 683

Db 499 GCGAGAGAGATGTCTAGTGGCCCCCTGAGAGAGAGCTGCGCCCTGCGCTGCGC 550

RESULT 15  
US-08-666-067-1

; Sequence 1, Application US/08666067

; Patent No. 5922842

; GENERAL INFORMATION:

; APPLICANT: Seedorf, Klaus

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT

; TITLE OF INVENTION: OF TKA-1 RELATED

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/666,067

; FILING DATE: June 13, 1996

; CLASSIFICATION: 530

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 60/005,421

; FILING DATE: October 13, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 220/157

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1642 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 97...1446

; US-08-666-067-1

Query Match 5.2%; Score 42.4; DB 2; Length 1642;  
Best Local Similarity 52.9%; Pred. No. 0.03; Mismatches 81; Indels 0; Gaps 0;  
Matches 91; Conservative 0;

QY 512 ctccagagcctctgagcagcatgagtgatgagcagcttgctaccactgccc 571

Db 379 CTCGCCCGCGGCGAGCTGTACCGAGAGATGGCCAGGAGGGCTCCACCGCGC 438

QY 572 cagcctctccagagactgagccagcctgagccctgagccctccacagtgactctc 631

Db 439 CAGGACCCCTGGAGCGCAAGCAGACTGGGCGACACACCGGCGACCAACAGCTCGAAGCT 498

QY 632 cagaagatgatgactctgctgctgaggaagactgagacactgctgctatgac 683

Db 499 GCGAGAGAGATGTCTAGTGGCCCCCTGAGAGAGAGCTGCGCCCTGCGCTGCGC 550

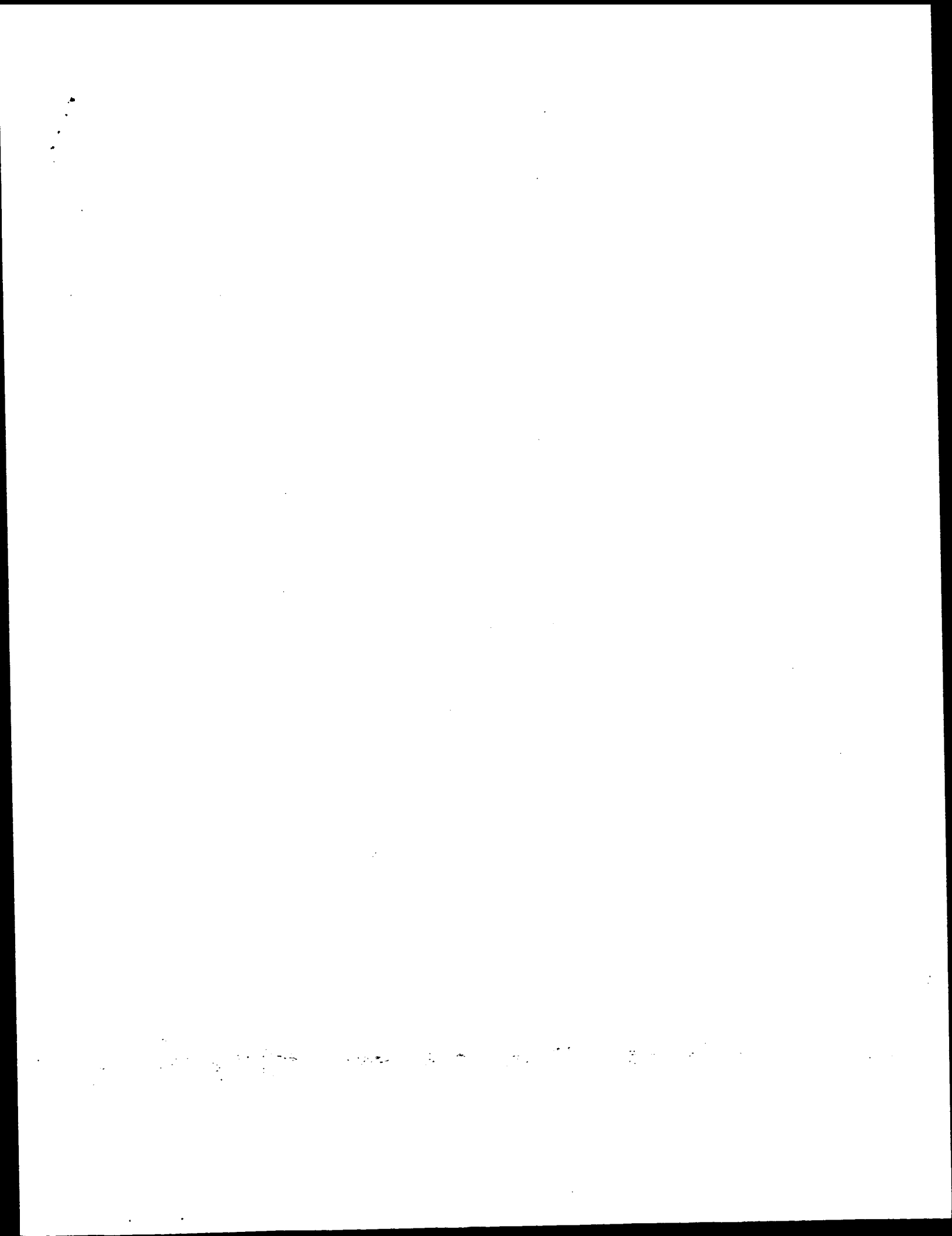
Mon Mar 18 11:23:41 2002

us-09-931-704-4.rn1

Page 11

Search completed: March 18, 2002, 10:25:14  
Job time: 2854 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:32:30 ; Search time 41.88 seconds  
(without alignments)  
409.247 Million cell updates/sec

Title: US-09-931-704-5  
Perfect score: 1224  
Sequence: 1 MDLRAGDSMGMLACTVLM.....KKKMPAASVTLHEAHGF 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.68.\*  
2: PIR1.\*  
3: PIR2.\*  
4: PIR3.\*  
5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118.5	9.7	203	2	JC4645
2	110.5	9.0	203	2	cardiotrophin-1 -
3	93.5	7.6	640	2	transferrin - Stre
4	92	7.5	332	2	hypothetical prote
5	91.5	7.5	201	2	cardiotrophin-1 -
6	90	7.4	195	2	ciliary neurotroph
7	90	7.4	733	2	hypothetical prote
8	89	7.3	733	2	conserved hypotet
9	86	7.0	1313	2	probable guanine n
10	85.5	7.0	389	2	probable cystathio
11	85.5	7.0	560	2	poly(3-hydroxyalka
12	85	6.9	520	2	extracellular solu
13	84	6.9	741	2	hypothetical prote
14	83	6.8	272	2	hypothetical prote
15	83	6.8	955	2	zinc finger protei
16	83	6.8	1561	2	zinc finger protei
17	80.5	6.6	1006	2	probable alpha-man
18	80	6.5	1179	2	TMV resistance pro
19	79.5	6.5	294	2	hupk protein - Rho
20	79.5	6.5	316	2	hypothetical prote
21	79.5	6.5	512	2	activin receptor S
22	79.5	6.5	542	2	hypothetical prote
23	79.5	6.5	644	2	anti-muellerian hor
24	79	6.5	1132	2	MHC class III hist
25	78.5	6.4	338	2	hypothetical prote
26	78.5	6.4	531	2	probable serine/th
27	78.5	6.4	571	2	probable potassium
28	78.5	6.4	1844	2	hypothetical prote
29	78	6.4	378	2	hypothetical prote

30	78	6.4	427	1	BRYC65
31	78	6.4	515	2	T37982
32	78	6.4	917	2	S40178
33	77.5	6.3	572	2	B72714
34	77.5	6.3	1182	2	I48378
35	77	6.3	322	2	T23891
36	77	6.3	411	2	F75439
37	77	6.3	1031	2	F83561
38	77	6.3	2261	1	A42548
39	76.5	6.2	368	2	T51111
40	76.5	6.2	395	2	I49575
41	76	6.2	232	2	E75547
42	76	6.2	508	2	A32225
43	76	6.2	720	2	T36293
44	76	6.2	812	2	T34180
45	76	6.2	891	2	H75507

#### ALIGNMENTS

##### RESULT 1

JC4645

C:Species: Rattus norvegicus (Norway rat)

C>Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 20-Jun-2000

C:Accession: JC4645

R;Shikawa, M.; Saito, Y.; Miyamoto, Y.; Kuwahara, K.; Ogawa, T.; Nakagawa, O.; Harad

Biochem. Biophys. Res. Commun. 219, 377-381, 1996

A:Title: cDNA cloning of rat cardiotrophin-1 (CT-1): Augmented expression of CT-1 gen

A:Reference number: JC4645; M0ID:96193659

A:Accession: JC4645

A:Molecule type: mRNA

A:Residues: 1-203 <TSH>

A:Cross-references: DDBJ:D78591; NID:g1256926; PIDD:BA11427 ; PID:g1256927

C:Genetics:

A:Gene: CT-1

C:Keywords: cardiac muscle; cytokine; heart

Query Match 9.7% Score 118.5; DB 2; Length 203;  
Best Local Similarity 27.5% Pred. No. 0.0012;  
Matches 49; Conservative 26; Mismatches 84; Indel 19; Gaps 6;

OY	40	IQKTYDITRYLEHQLRSIAGTYLNYLGPPEPNEPDPNPRLL---GAETLPK...VNLKFWRS 96
DB	27	IRQTHNLARLLTKYADQLLEYVOQGEPFGLPGFSPRLPLAGLSGPAF...NLLPV--- 83
OY	97	LNDRLRLTONYENYSHLCYRGLNRQAA-----TAEIRSLAHCFTSLGCIICSIAGVM 151
DB	84	-SERLR--QDAALASALPALDDAVARRQAEINPRAPRLRLSEDAARQVRALGAATETVL 140
OY	152	ATLGY----PLPQPLPGTEPAWAPGPAHSDFLOKMDDEFWLKELOTWLR*NDPNNRL 205
DB	141	AALGAAGPVEPY-ATSALETSSAGVSAKVLGLHVCGLYGEWVSRT...DLGOL 197

##### RESULT 2

cardiotrophin-1 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999

C:Accession: I49153

R;Pennica, D.; King, K.L.; Shaw, K.J.; Luis, E.; Rullman, J.; Dharbone, W.

Proc. Natl. Acad. Sci. U.S.A. 92, 1142-1146, 1995

A:Title: Expression cloning of cardiotrophin 1, a cytokine that induces cardiac myocy

A:Reference number: I49153; M0ID:95166785

A:Accession: I49153

A>Status: nucleic acid sequence not shown; translated from cDNA

A:Molecule type: mRNA

A:Residues: 1-203 <RES>

A:Cross-references: EMBL:018366; NID:g710331; PIDD:AAC5217 ; PID:g710332

C:Genetics:

A:Gene: cctf1  
C:Keywords: cytokine; glycoprotein  
F:164/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Query Match

Best Local Similarity 9.08; Score 110.5; DB 2; Length 203;  
Matches 49; Conservative 23; Mismatches 86; Indels 21; Gaps 6;

40 IQKTYDRLRYLEHQLRSLAGTYLNYLGPENPDPPRL---GAETLPRTATVLEWRS 96  
Db IROTHNARLRTYAEQLLEEVYQOQCEPGLPEFSPRLPLAGLSGAPSHAGLPV--- 83  
QY 97 LNDRLRLTONYEAYSHLLCYLRGLNRQA-----TALRSLSAHFCTSLGGLGSGVM 151  
Db 84 -SERLR--QDAALSVLPALLDAVRRQALNRPRLRLSLDEAQAQVALGAAYETVL 140  
QY 152 ATGTYLPLPQLPCTEPA-----WAPGPAHSDFLQKMDPWLKELQTLWRSKADENRL 205  
Db 141 ALGAAARG--PEPEPVYATLETANSTAGIFSAKYLGFEHVCGLYGEWVSRTEGDLGOL 197

## RESULT 3

T34916

transferase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 02-Sep-2000

C:Accession: T34916

R:Olliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

A:Submitted to the EMBL Data Library, January 1998

A:Reference number: 221558

A:Accession: T34916

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-640 &lt;OLI&gt;

A:Cross-references: EMBL:AL021409; PIDN:CAA16181.1; GSPDB:GN00070; SCOEDB:SC3F7.10

A:Experimental source: strain A3(2)

A:Genetics:

A:Gene: SCOEDB:SC3F7.10

A:Superfamily: glycine C-acetyltransferase homology

F:287-624/Domain: glycine C-acetyltransferase homology &lt;GCA&gt;

## Query Match

Best Local Similarity 7.6%; Score 93.5; DB 2; Length 640;  
Matches 52; Conservative 12; Mismatches 63; Indels 57; Gaps 9;

22 LPAYPALNRGDDPGPSIQKTYDRLRYLEHQL---RSLAGTYLNYLGPENPDPPR 78  
Db 100 LPAYPA-EGTGDPTVEAVVAAMRQYQRHQLGDADLEG-----E 140  
QY 79 LGATTLPRATVNLFWRSRLNDRRLTONYEAYSHLLCYLRGLNRQAATAELRSLAHFCT 138  
Db 141 LGVDSVLTVAFA-----TERLGLT-----GAAPDAAGATTLRALAA--D 179  
QY 139 SLGGLLSIAGVMAATLGYPRLPGTEPAPAPGPAHSDFLQKMDPWLKELQTLWRS 198  
Db 180 ALRGL-----VAAAPGTAPEAPATGAA-APAPGRSGNAP-----ACGAGCGDHRS 225

QY 199 AKDF 202  
Db 226 MKDF 229

## RESULT 4

G86182

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: G86182  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khoo, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, S.; Maiti, R.; Marzita,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Soultz, A.M.; Sun, H.; Tallo,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *A. thaliana*.

A:Reference number: A86141; MUID:21016719

A:Accession: G86182

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-332 &lt;STO&gt;

A:Cross-references: GB:AE005172; NID:g7211973; PIDN:AAF0444; SVD:GN00141

A:Map position: 1

## Query Match

Best Local Similarity 7.5%; Score 92; DB 2; Length 332;  
Matches 57; Conservative 24; Mismatches 73; Indels 10; Gaps 12;

QY 23 PAVPALNRGDDPGPSIQKTYDRLRYLEHQLRSLAGTYLNY----- 64  
Db 113 PSYTAGNLGYP-PRPSF--TDPGEYEQOMESLIQOFIRENPOLKPL-GLGSPVG 169  
QY 65 LGPPFNPDPNPRLGAETLPRTATVLEWRSRLNDRRLTONYEAYSHLLCYLRGLNRQA 124  
Db 170 LGPIRASPQFLQPRVA-----PPPTSLIDTSRNRRKRSK-----DGLAVVRG--RKV 215  
QY 125 ATAELRSL-----AHFCTSLGGLLSIAGVMAATLGYPRLPGTEPAPA 170  
Db 216 RITEGSSSLYSLSGRSMKNGAH-----GIQOPRSSIMK---PLKPLPVVHVTETSV 266  
QY 171 PGPAHSDFLQKMDPWLKELQTLWRSKADENRLKKKK 210  
Db 267 DDPDEESADEKDEEAVAKOL-----SEKDL--LKRHI 298

## RESULT 5

G02312

cardiotrophin-1 - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998

C:Accession: G02312

R:Wood, W.I.

A:Submitted to the EMBL Data Library, December 1995

A:Reference number: H01035

A:Accession: G02312

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-201 &lt;WOO&gt;

A:Cross-references: EMBL:U43030; NID:g1151149; PID:g1151150

A:Genetics:

A:Gene: GDB:CFP1; CT-1

A:Cross-references: GDB:567078

A:Map position: 1p22-1p22

## Query Match

Best Local Similarity 7.5%; Score 91.5; DB 2; Length 332;  
Matches 45; Conservative 23; Mismatches 92; Indels 15; Gaps 5;

QY 40 IQKTYDRLRYLEHQLRSLAGTYLNYLGPENPDPPRLGAETLPRTATVLEWRS 99  
Db 27 IROTHNARLRTYAEQLLEEVYQOQCEPGLPEFSPRLPLAGLSGAPSHAGLPV--- 83  
QY 100 RLRLTONYEAYSHLLCYLRGLNRQA-----TALRSLSAHFCTSLGGLGSGVM 154  
Db 86 RLRL--DAALAAALPRLDVAVCRRQALNRPRLRLRLLEDAQAQVALGAAYETVL 143  
QY 155 GY-----PLPPLPTEPAPAPGPAHSDFLQKMDPWLKELQTLWRSKADENRL 205  
Db 144 GAANRGPRAAPPAATASA---ASATGVPAKVLGLRVCGLYREWLSTEGH--NLC 195

RESULT 6  
JH0680  
ciliary neurotrophic factor - chicken

N:Alternate names: growth-promoting activity protein

C:Species: Gallus gallus (chicken)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jul-2000

C:Accession: JH0680; PQ0057

C:Length: 1,045-1,053, 1992

A:Title: Cloning, expression during development, and evidence for release of a trophic factor

A:Reference number: JH0680; MUID:92304573

A:Accession: JH0680

A:Molecule type: mRNA

A:Residues: 1-195 <LEU>

A:Cross-references: GB:M80827; NID:g211822; PIDN:AAA48784.1; PID:g211823

A:Experimental source: eye

R:Eckenstein, F.P.; Esch, F.; Holbert, T.; Blacher, R.W.; Nishi, R.

Neuron 4, 623-631, 1990

A:Title: Purification and characterization of a trophic factor for embryonic peripheral

A:Reference number: PQ0057; MUID:90211978

A:Accession: PQ0057

A:Molecule type: protein

A:Residues: 155-166, X, 168-175 <ECK>

A:Experimental source: sciatic nerves

C:Comment: This is a neurotrophic protein.

C:Superfamily: ciliary neurotrophic factor

C:Keywords: growth factor

Query Match 7.4%; Score 90; DB 2; Length 195;

Best Local Similarity 27.7%; Pred. No. 0.48;

Matches 52; Conservative 21; Mismatches 83; Indels 32; Gaps 10;

QY 46 LTRYLEHQLSLAGTYLNTLGPFPNPPRLGAEITLPRATVNLVWRSLSLNDRLRLTQ 105

DB 23 LARKMSDVIDLDLDIVYERG-----LDASISVAANDGVPTAAV--ERNAEQTGTORLDD 75

QY 106 N---YEAYSHLTCYLRGLNROA---ATAELRSLA-----HFCSTSLGSLGSIAGVMA 152

DB 76 NLAYRFRFTLLQMLEEGRELLGDTDAELGPAALAMLVSAFVYHLEEL-----ELE 130

QY 153 TLGPIPLPQPLGTEPPAPAHSDPLQKMDPWLKELQTLWLRSAKDKENRLKKMKOP 212

DB 131 SRGAPAE---GSEPP-AP-PRLSLFEQKLRGLRLELQAWVRSVRLRLSK--HGP 183

QY 213 AASVTLLH 220

DB 184 GSGAALGL 191

RESULT 7  
F82965  
hypothetical protein PA5441 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: F82965

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

adman, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337

A:Accession: F82965

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-733 <STO>

A:Cross-references: GB:AE004957; GB:AE004091; NID:g9951770; PIDN:AAG08826.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA5441

Query Match 7.4%; Score 90; DB 2; Length 733;

Best Local Similarity 23.9%; Pred. No. 2.5;

Matches 48; Conservative 24; Mismatches 81; Indels 48; Gaps 8;

QY 23 PAVPALNRGTGDPGPGSTIQKTYDITRLEHQLSLAGTYLNTLGPFPNPPRLGAEITLPRATVNLVWRSLSLNDRLRLTQ 82

DB 102 PALPAAAEATPPAPAPSPALAEAR-----OMGAE 132

QY 83 TLPRATVNLVWRSLSLNDRLRLTQNEYAYSHLTCYLRGLNROATAELE--HFCSTSLG 141

DB 133 ALPEKINGAEFFEGEGSRCSNDQSA---LAFRLROY-RDAGLGEATKALAN---SR 184

QY 142 GLGSLGSIAGVMAATGTYPLPQPLGTEPPAPAHSDPLQKMDPWLKELQTLWLRSAKDKENRLKKMKOP 201

DB 185 DLGACGMEDEBELGCVLAQ--GVES--AAGKAFATYLEAANFYSGR--FDEAEG 224

QY 202 FNRLKKMKOPPAASVTLLHLEA 222

DB 235 FKALQDVSQLPKETALYLQA 255

RESULT 8  
D83588  
conserved hypothetical protein PA0454 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: D83588

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

adman, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: D83588

A:Molecule type: protein

A:Residues: 1-733 <STO>

A:Cross-references: GB:AE004483; GB:AE004091; NID:g994631; PIDN:AAG03843.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0454

C:Superfamily: hypothetical protein H11680

Query Match 7.3%; Score 89; DB 2; Length 733;

Best Local Similarity 25.8%; Pred. No. 3.1;

Matches 57; Conservative 26; Mismatches 64; Indels 74; Gaps 13;

QY 28 LNRGTGDPGPGSTIQKTYDITRLE-----HQLSLAGTYLNTLGPFPNPPRLGAEITLPRATVNLVWRSLSLNDRLRLTQ 77

DB 230 LNRGLHGRGPKV-----SRYLKLYFIADQVHERASSSHYNNRLAEATFPHSDVLFRCQ 283

QY 78 RL-----GAETLPRATVNLVWRSLSLNDRLRLTQNEYAYSHLTCYLRGLNROATAELE--HFCSTSLG 131

DB 284 RLNDQGGKACQALARA-----IRLRQPPD-YA-----DL-LALEDLQA 320

QY 132 SLAHF-----CTSLQGI--LGSITAGVMAATGTYPLPQPLGTEPPAPAHSDPLQKMDPWLKELQTLWLRSAKDKENRLKKMKOP 184

DB 321 SLHLRQGSNPAWKGLRSLGALANLITLDRKL-----AGASNIATADDEQ 368

QY 185 FWLKEQLQTLWLRSAKDKENRLKKMKOPPAASVTLLHLEAG 224

DB 369 SALDRSP-----RSLKDAFERLRQQLTP-----TSLFLRHG 400

RESULT 9  
T38943  
probable guanine nucleotide binding protein - fission yeast - Schizosaccharomyces pombe

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T38943

R:Skellon, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.; Wood, V.

submitted to the EMBL Data Library, May 1997

A:Reference number: 221819  
A:Accession: T38943  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1313 <SKED>  
A:Cross-references: EMBL:295396; PIDN:CA08769.1; GSPDB:GN00066; SPDB:SPAC57A7.11  
A:Experimental source: strain 972h; cosmid c57A7  
C:Genetics:  
A:Gene: SPDB:SPAC57A7.11  
A:Map position: 1

## Query Match

Best Local Similarity 7.0%; Score 86; DB 2; Length 1313;  
Matches 58; Conservative 31; Mismatches 84; Indels 88; Gaps 12;

OY 15 LCTVHLHPLAV-PALNR-----TGDGPGPST-----OKTYDLTRYLEHQLRS 56  
DB 811 LAFLLQHLPLHKAASLSDKDTGNSVTSDEKPHFVPSVSENKILNRSFSLRSKGLALS 870  
OY 57 LAG-----TYLNL-----GPP-FNEPPDNPPRLGAETLPRATYNL 91  
DB 871 LAGSDRASELLSLNGENKPAESNLNHLTSAKYGPFAFELEYQ-----SELDMLTSLYF 926  
OY 92 EWRSLNDLRLTONYEAYSHLLCYLRGL--NRQAATAELRSLAHCTS-----LOG 142  
DB 927 DMSRRYFTEPQMRPNEDDEPGSICYNORLMRRNREKLIYRTRPLAEYSTNGRMNQOLMT 986  
OY 143 ILGSIAGVATIGYPLPPLPCTEPAMAPGPAHSDFLQKMDFWLKE---LOTWLMKSA 199  
DB 987 FNTTIA-----PRKLMFHOFEDQLTLTGDKDIIQVMDWRN 1022  
OY 200 KDFNLRKKMKOPPAATVTLH 220  
DB 1023 RCLNSFK-----TSASATTVN 1038

## RESULT 10

F72511

Probable cystathionine gamma-synthase APE2068 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: F72511  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6: 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
A:Reference number: A72450; MUID:99310339  
A:Accession: F72511  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-389 <KAW>  
A:Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BA081078.1; PID:95105766  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2068  
C:Superfamily: O-succinylhomoserine (thiol)-lyase

## Query Match

Best Local Similarity 7.0%; Score 85.5; DB 2; Length 389;  
Matches 47; Conservative 20; Mismatches 77; Indels 61; Gaps 8;

OY 43 TYDLTRYLEHQLRSIAGVATIGYPLPPLPCTEPAMAPGPAHSDFLQKMDFWLKELOTWLMRS 97  
DB 107 TYGSTSLLEMLSSITIGIEVRLAGPFWED-----LLDLYCWADLIVES 150  
OY 98 --NDRLRLTONYEAYSHLLCYLRGLNRQAATAELR-----RSL---AHFCTS 139  
DB 151 MANPTRLVPP-----LSGIYRAGSGCVRVVVNTFATPIAYNPLRGAIY--S 197  
OY 148 LGLLGIAGVATIGYPLPPLPCTEPAMAPGPAHSDFLQKMDFWLKELOTWLMRS 198

DB 198 LESLTKYIAGHNDVYGGSLSGREEDLEPIMMKRIITQMPIDAY-----LAWRG 249  
OY 199 AKDFNLRKKMKOPPAATVTLHLEAH 223  
DB 250 MTKLAREPAOSRAVEAWELESH 274

## RESULT 11

C38604

poly(3-hydroxyalkanoate) polymerase (EC 2.7.7.-) 3 - Pseudomonas oleovorans  
C:Species: Pseudomonas oleovorans  
C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 08-Sep-2000  
C:Accession: C38604  
R:Huisman, G.W.; Monink, E.; Meima, R.; Kazemier, B.; Terpstra, P.; Wiltolt, B. J. Biol. Chem. 266, 2191-2198, 1991  
A:Title: Metabolism of poly(3-hydroxyalkanoates) (PHAs) by Pseudomonas oleovorans. Id  
A:Reference number: A38604; MUID:91115830  
A:Accession: C38604  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-560 <HUI>  
A:Cross-references: GB:M58445; NID:9151441; PIDN:AA25941; PID:9151444  
C:Superfamily: poly(3-hydroxyalkanoic acid) synthase phb  
C:Keywords: nucleotidyltransferase

## Query Match

Best Local Similarity 7.0%; Score 85.5; DB 2; Length 46;  
Matches 38; Conservative 16; Mismatches 43; Indels 29; Gaps 7;

OY 77 PRGAETLPRATVLEWRSLNDLRLTONYEAYSHLLCYLRGLNRQAATAELR---A 134  
DB 5 PAKSTPLPAPTSMMVQ-----NALIGLRG-----DLISLRNVRSGS---LHETHTA 50  
OY 135 HFCTSLGGLSIAGVATIGYPLPPLPCTEPAMAPGPAHSDFLQKMDFWLKE 190  
DB 51 HHLALGGQLGRV-----ILGDTPLQPNRPDRFSDPTSONPFRKTYDA--YLAHQ 101  
OY 191 LOTWLM 196  
DB 102 KQTRLW 107

## RESULT 12

B75332

extracellular solute-binding protein, family 5 - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: B75332  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J. S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: B75332  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-520 <WHI>  
A:Cross-references: GB:AE002034; GB:AE000513; NID:96459742; NID:NAF11508.1; PID:9645  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1955  
A:Map position: 1  
C:Superfamily: dipeptide transport protein

## Query Match

Best Local Similarity 6.9%; Score 85; DB 2; Length 52;  
Matches 40; Conservative 20; Mismatches 54; Indels 5; Gaps 7;

OY 19 LWHLPAYPALNRTGDPGPGPSIQ-----KTYDLTRYLEHQLRSIAGVATIGYPLN 63



Db 363 LMY-----PSISRYVLPKPKVFAEIAADLGAIGIKVNLKTEDWPKYLEDDBRAGFGQWLY 418  
 QY 64 YLGPPEEPD-----FNPRPGAETLPRAVNLVWRSIN-DRLRLTON 106  
 Db 419 GMGGDYNDPDPNFYSARYGDKGSDDIGFNG-----NILNTLNKGRALTOA 464  
 QY 107 YEAYSHLLCYLRLGNROAATLRLSLAHECTSLG-----LLGSAGVMTLG 155  
 Db 465 OKA-----PYKOLHELTYNANVRLPLVHSTAAARTYKGMTGPHSLIVG 511

## RESULT 13

A83271

hypothetical protein PA2984 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: A83271

R:Stover, C.K.; Pham, X.Q.; Ertyn, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lim, Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: A83271  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-741 <STO>  
 A:Cross-references: GB:AE004724; GB:AE004091; NID:99949083; PIDN:AA06372.1; GSPDB:GNO01  
 C:Genetics:  
 A:Experimental source: strain PA01  
 A:Gene: PA2984

Query Match 6.9%; Score 84; DB 2; Length 741;  
 Best Local Similarity 26.2%; Pred. No. 9;  
 Matches 64; Conservative 13; Mismatches 81; Indels 86; Gaps 13;

QY 5 AGDSW-----GMLACTGVNHLPAVPAVLRNRTGDPGPGPSIOKTYDLTRYLEHQLSLAGT 60  
 Db 447 AAAASVALGMLGALC---WLAPA-----GLP-----LRALGAA 477  
 QY 61 YLNLGPPFNEPDPFNPRLGAETLPRAVNLVWRSIN-DRLRLTONYEASHLLCYLRG 119  
 Db 478 LL-----LPALLPSPPEVWEG-----RAEVRVLDYGGLAVLVKRE-----HVLVDSG 522  
 QY 120 LNRQA-----TAELRSLAHFCTSLGGLGSLAGVMTLGYP-----PQ 160  
 Db 523 A-RQGAADMGERRVVPVLRSLDLRR-----LDGILLSHADNDHAGAPTVASRPPV 573  
 QY 161 PLPTEPAPAPGPAHSDFLQKMDFFMLKELQTMWMSAKDFNRLKKMKPPASVTLHL 220  
 Db 574 WLVSGEPAARLPPLFLFADSCDERSMSMDGVVEQWAMAQADSN-----DRSCVLRV 624  
 QY 221 EAHG 224  
 Db 625 EADG 628

## RESULT 14

T35231

hypothetical protein SC5C7.24c SC5C7.24c - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T35231

R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, September 1998  
 A:Reference number: 221572  
 A:Accession: T35231  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-272 <SEE>  
 A:Cross-references: EMBL:AL031515; PIDN:CAA20636.1; GSPDB:GN00070; SCOEDB:SC5C7.24c

A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC5C7.24c

Query Match 6.8%; Score 83; DB 2; Length 741;  
 Best Local Similarity 23.9%; Pred. No. 3.2;  
 Matches 38; Conservative 15; Mismatches 50; Indels 56; Gaps 7;

QY 23 PAVPALN-----RIGDPGPGPSIOKTYDLTRYLEHQLSLAGTYLNTY-----FN----- 70  
 Db 95 PALGALREGMERLIRGYEPYPAVAVATYRVLAAANGIAMMDGVASHLLI-----NATRLT 154  
 QY 71 -EPDPNPRLGAETLPRAVNLVWRSIN-DRLRLTONYEASHLLCYLRG 129  
 Db 155 LHPDGLAPRIR-----NLREWR-----GHL-----EOM 178  
 QY 130 RSLA-HFCTSLGGLGSLAGVMTLGYPPLPDPTEP 167  
 Db 179 EROIALHRSRPLRELYDEVA-----AYPVESVGAEP 211

## RESULT 15

T00247

zinc finger protein w12 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 05-Nov-1999  
 C:Accession: T00247

R:Matsmoto, K.; Ishii, N.; Yoshida, S.; Shiosaka, S.; Waukada, T.; Tohyama, M. submitted to the EMBL Data Library, March 1998  
 A:Description: Molecular cloning and distinct developmental expression pattern of spl  
 A:Reference number: Z14130  
 A:Accession: T00247  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-955 <MAT>  
 A:Cross-references: EMBL:AB012266; NID:d1227741; PIDN:BA03279.1; PID:d1033757  
 A:Experimental source: brain  
 A:Genetics:  
 A:Gene: w12

Query Match 6.8%; Score 83; DB 2; Length 955;  
 Best Local Similarity 27.0%; Pred. No. 15;  
 Matches 34; Conservative 20; Mismatches 48; Indels 41; Gaps 7;

QY 22 LPAVPAVNRGDPGPGPSIOKTYDLTRYLEHQLSLAG---TYLNTYLP-----PFPNPD 73  
 Db 615 LPLSPPLASRPKPGAGPT-----QVPR--ELSLPITGSKPSAASYGLVA-----PFLQEDR 667  
 QY 74 FNPRPGAETLPRAVNLVWRSIN-DRLRLTONYEASHLLCYLRG-----RQGAATAELR 130  
 Db 668 FLPAEVAKAKITYTETLPKKA-KTLHER-----TSHSSTEACCELGIVPEWKKALASHAR 721  
 QY 131 RSLAHF 136  
 Db 722 AHLROF 727

Search completed: March 18, 2002, 09:32:32  
 Job time: 218 sec

Mon Mar 18 11:23:42 2002

us-09-931-704-5.rpt

---



QY 97 LNDRLRLTONYEAAYSHLLCYLRGLNROA-----TAEIIRSLAHFCTSLQGLGSIAGVM 151  
 Db 84 -SERLR--ODAAALSLPALLDAVRRROAELNRPAPRLRSLEDAARQVRAALGAAYETVL 140  
 QY 152 ATGCGY-----PLPQPLPCTEPAPAPGPAHSDFLQKMDPFWLKLKELQTLWLRSAKDKNRL 205  
 Db 141 AALGAARAGVPEPV-ATSLALFTSNAGVSAKVLGLHVCLYXGEMVSRTEGDLGQL 197

## RESULT 2

CTFL\_MOUSE  
 ID CTFL\_MOUSE STANDARD: PRT: 203 AA.  
 AC 060753;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE 15-JUL-1998 (Rel. 36, Last annotation update)  
 GN CARDIOTROPIN-1 (CT-1).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9516785; PubMed=7862649;  
 RA Pennica D., King K.L., Shaw K.J., Luis E., Rullamas J., Luo S.-M.,  
 RA Darbonne W.C., Knutson D.S., Yen R., Chien K.R., Baker J.B.,  
 RA Wood W.I.;  
 RT "Expression cloning of cardiostrophin 1, a cytokine that induces  
 RT cardiac myocyte hypertrophy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:1142-1146(1995).  
 CC -1- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO  
 CC AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF  
 CC RECEPTOR)/GP 130 RECEPTOR COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,  
 CC LIVER, LUNG AND KIDNEY. LOWER LEVELS IN TESTIS AND BRAIN. NO  
 CC EXPRESSION IN SPLEEN.  
 CC -1- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; U18366; AAC52173.1;  
 DR MGD; MGI:105115; Ctf1.  
 KW Cytokine.  
 SQ SEQUENCE 203 AA: 21509 MW: 883D414A0B3B232F CRC64;

Query Match 9.0%; Score 110.5; DB 1; Length 203;  
 Best Local Similarity 27.4%; Pred. No. 0.0023;  
 Matches 49; Conservative 23; Mismatches 86; Indels 21; Gaps 6;

QY 40 IOKTYDTRLEHQLSLAGTYLNTLGPFPNPPDNPRL--GAETLPRAVTLNFWRS 96  
 Db 27 IROTHSLAHLLTKRYAQQLQEVYVQOGDFPGFPSPRLPLAGLSGAPSHAGLPV-- 83  
 QY 97 LNDRLRLTONYEAAYSHLLCYLRGLNROA-----TAEIIRSLAHFCTSLQGLGSIAGVM 151  
 Db 84 -SERLR--ODAAALSLPALLDAVRRROAELNRPAPRLRSLEDAARQVRAALGAAYETVL 140  
 QY 152 ATGCGY-----PLPQPLPCTEPAPAPGPAHSDFLQKMDPFWLKLKELQTLWLRSAKDKNRL 205  
 Db 141 AALGAARAGVPEPV-ATSLALFTSNAGVSAKVLGLHVCLYXGEMVSRTEGDLGQL 197

## RESULT 3

CTFL\_HUMAN  
 ID CTFL\_HUMAN STANDARD: PRT: 195 AA.  
 AC 002011;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)

ID CTFL\_HUMAN STANDARD: PRT: 201 AA.  
 AC 016619;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE 20-AUG-2001 (Rel. 40, Last annotation update)  
 GN CARDIOTROPIN-1 (CT-1).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homiidae;  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Heart;  
 RX MEDLINE=96429882; PubMed=8833032;  
 RA Pennica D., Swanson T.A., Shaw K.J., Kuang W.-J., Gronowicz  
 RA Beatty B.G., Wood W.I.;  
 RT "Human cardiostrophin-1: protein and gene structure, biologic and  
 RT binding activities, and chromosomal localization.";  
 RL Cytokine 8:183-189(1996).  
 CC -1- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO  
 CC AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF  
 CC RECEPTOR)/GP 130 RECEPTOR COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, SKELETAL MUSCLE,  
 CC THYMUS, TESTIS AND OVARY. LOWER LEVELS IN LUNG, KIDNEY, PANCREAS,  
 CC BRAIN, PLACENTA, LIVER, SPLEEN, COLON OR PERIPHERAL BLOOD  
 CC LEUCOCYTES.  
 CC -1- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; U43033; AAD12173.1;  
 DR EMBL; U43031; AAD12173.1; JOINED.  
 DR EMBL; U43032; AAD12173.1; JOINED.  
 DR HSSP; P04002; AA85229.1;  
 DR MIM; 600435;  
 KW Cytokine.  
 SQ SEQUENCE 201 AA: 21227 MW: 0235A7B5745F675F CRC64;

Query Match 7.5%; Score 91.5; DB 1; Length 195;  
 Best Local Similarity 25.7%; Pred. No. 0.14;  
 Matches 45; Conservative 23; Mismatches 92; Indels 15; Gaps 5;

QY 40 IOKTYDTRLEHQLSLAGTYLNTLGPFPNPPDNPRLGAETLPRAVTLNFWRS 96  
 Db 27 IROTHSLAHLLTKRYAQQLQEVYVQOGDFPGFPSPRLPLAGLSGAPSHAGLPV-- 83  
 QY 100 RLRLTONYEAAYSHLLCYLRGLNROA-----TAEIIRSLAHFCTSLQGLGSIAGVM 151  
 Db 86 RLRL--DAALALPALLDAVRRROAELNRPAPRLRSLEDAARQVRAALGAAYETVL 140  
 QY 155 GY-----PLPQPLPCTEPAPAPGPAHSDFLQKMDPFWLKLKELQTLWLRSAKDKNRL 205  
 Db 144 GAANRGPRAEPPAATASA--ASATGVFPAKVLGLRVCGYREWLSTEGDLGQL 195

## RESULT 4

CTFL\_CHICK  
 ID CTFL\_CHICK STANDARD: PRT: 195 AA.  
 AC 002011;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE CILIARY NEUROTROPHIC FACTOR (CNTF) (GROWTH PROMOTING ACTIVITY) (GPA).  
 GN CNF OR GPA.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NCBI\_TaxID=9031;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92304573; PubMed=1610564;  
 RA Leung D.W., Parent A.S., Cachianes G., Lee A.L., Nikolic K.,  
 RA Esch F., Coulombe J.N., Blacher R.W., Eckenstein F.P., Nishi R.;  
 RT "Cloning, expression during development, and evidence for release of  
 a trophic factor for ciliary ganglion neurons."  
 RL Neuron 8:1045-1053(1992).  
 CC -1- FUNCTION: CNTF IS A SURVIVAL FACTOR FOR VARIOUS NEURONAL CELL  
 TYPES. SEEMS TO PREVENT THE DEGENERATION OF MOTOR AXONS AFTER  
 AXONOMY.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM.  
 CC -1- SIMILARITY: BELONGS TO THE CNTF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M80827; AAA48784.1; -  
 CC DR EMBL: 295396; CAB08769.1; -  
 CC DR InterPro: IPR001680; WD40.  
 CC DR HSP: P26441; ICNT.  
 CC DR InterPro: IPR000151; CNTF.  
 CC DR Pfam: PF01110; CNTF; 1.  
 CC DR ProDom: PD011041; CNTF; 1.  
 CC Growth factor; Neurone.  
 CC KW SEQUENCE 195 AA; 21330 MW; FEA076949DB34AC5 CRC64;  
 SQ

Query Match 7.4%; Score 90; DB 1; Length 195;  
 Best Local Similarity 27.7%; Pred. No. 0.19;  
 Matches 52; Conservative 21; Mismatches 83; Indels 32; Gaps 10;

QY 46 LTRYEHQRLSAGTYLWLPPEPPNPRLGAEPLPRATVNLVWRSINDRLRLTQ 105  
 Db 23 LARKKRSYVTLDDIYERQG-----IDASISVAADVPTAAV--ERWAEQTGTORILD 75  
 QY 106 N---YEASHLICYLRLGLNROA---ATAELRRSLA-----HCTSLQGLISAGYMA 152  
 Db 76 NLAAYRAFRLLAQMLEQRELLDDTDALGLPALAAMLLQSAFYHLEELL-----ELE 130  
 QY 153 TLGYPLQPLRGTEPPANAPGAHSDFLQKMDFWLLKELQTLWMSAKDFNRKKMKOP 212  
 Db 131 SRGPAEE---GSEPP-AP-PRLSLFQKRLKGLHYRLRLAQMVAVSVDLQLSK--HGP 183  
 QY 213 AASVTLHL 220  
 Db 184 GSGAALGL 191

RESULT 5  
 MIPL\_SCHPO STANDARD; PRT; 1313 AA.  
 AC P87141;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE WD-REPEAT PROTEIN MIPL.  
 GN MIPL OR SPAC57A7.11.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.  
 OC NCBI\_TaxID=4896;  
 RN (1)  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=20115869; PubMed=10648609;  
 RA Shinozaki-Yabana S., Watanabe Y., Yamamoto M.;  
 RT "Novel WD-repeat protein Miplp facilitates function of the meiotic  
 RT regulator Mei2p in fission yeast."  
 RL Mol. Cell. Biol. 20:1234-1242(2000).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Skelton J., Churcher C.M., Wood V., Barrell B.G., Rajaniemi M.A.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBD database.  
 CC -1- FUNCTION: BINDS TO AND FACILITATES THE FUNCTIONING OF THE MEIOTIC  
 CC REGULATOR MEI2. MAY ALSO BE INVOLVED IN CONUGATION BY INTERACTING  
 CC WITH STE11. ESSENTIAL FOR CELL GROWTH.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB032552; BAA84585.1; -  
 CC DR EMBL: 295396; CAB08769.1; -  
 CC DR InterPro: IPR001680; WD40.  
 CC DR Pfam: PRO0400; WD40; 7.  
 CC DR PRINTS: PRO00320; GPROTEINBRPT.  
 CC DR SMART: SM00320; WD40; 5.  
 CC DR PROSITE: PS00678; WD-REPEATS\_1; 1.  
 CC DR PROSITE: PS50082; WD-REPEATS\_2; 2.  
 CC DR PROSITE: PS50294; WD-REPEATS\_REGION; 1.  
 CC KW Meiosis; WD repeat; Repeat.  
 CC FT REPEAT 986 1029 WD 1.  
 CC FT REPEAT 1033 1074 WD 2.  
 CC FT REPEAT 1087 1126 WD 3.  
 CC FT REPEAT 1130 1170 WD 4.  
 CC FT REPEAT 1176 1216 WD 5.  
 CC FT REPEAT 1219 1259 WD 6.  
 CC FT REPEAT 1268 1308 WD 7.  
 CC SEQUENCE 1313 AA; 148533 MW; C71B663B0171E7A4 CRC64;  
 SQ

Query Match 7.0%; Score 86; DB 1; Length 1313;  
 Best Local Similarity 22.2%; Pred. No. 4.9;  
 Matches 58; Conservative 31; Mismatches 84; Indels 48; Gaps 12;

QY 15 ICTYVIMHLPAY--PALNR-----TGDPGPGPSI-----OKTYDVLLEHQLRS 56  
 Db 811 LAFLLQHPALHKAASLSDTDTSVSDPKPPEVPSVSENKILKRSI--LSKGLALS 870  
 QY 57 LAG-----TYLNYL-----GPP-FNEPDPNPRLGAEPLPRATVNL 91  
 Db 871 LAGSDRASELLSINGENKPAESNLNHLTSAKVPGPAPNELEYQ---SEIIMPLTSLF 926  
 QY 92 EWMRSINDRLRLTONYEAVSHLLCYLRLG---NRQATAELRSLAHFCT---LQG 142  
 Db 927 DMSRRYFTEPQMRNEDDEPSICYNQRLMRNRNEKLIYTRPLAFV--LNKMQQLMT 986  
 QY 143 LLGSLAGVMTLGYPLPPLRGTEPPANAPGAHSDFLQKMDFWLLKEL--TWLWRS 199  
 Db 987 FNNRTA-----PRKLTFQEFQDLITLGHKDI--VWDWBRN 1022  
 QY 200 KDNRLKKMKOPPAASYTLHL 220  
 Db 1023 RCINSFK-----TSASATTVV 1038





FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1 216 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 377 377 A > R (IN REF. 1).  
 FT CONFLICT 603 603 MISSING (IN REF. 1).  
 SQ SEQUENCE 830 AA; 92997 MW; 62EACAE9A2D0C698B CRC64;

Query Match 6.5%; Score 79.5; DB 1; Length 830;  
 Best Local Similarity 24.3%; Pred. No. 11;  
 Matches 45; Conservative 25; Mismatches 66; Indels 49; Gaps 9;

QY 40 IOKTYDLTRYLEHQLSLAGTYLNYLGPENPPPPRCAETLPATVLEWRSIND 99  
 DB 61 LEKFE---TFLOEEVRR-AGL---VLPPPKGRPLPAPPR-----D 93  
 QY 100 RLRTQNEAVSHLCTVLRGLNROAATRLRSIAHCTSLGGLGSIAGVW---ATLGY 156  
 DB 94 LRLQETRLAQLQDVRG--NQQLRLQHLQDLHAVALRQGHPEPOLAAHTGASERT 152  
 QY 157 PLPQPLPTEPAPAPGPAHSD---FLOKMDFWLKLQGLWMSAK---DENRLKK 207  
 DB 153 PLQO-----APGQHQLRVNFGAVGEPHKAPELRLMRACGFLIASFRELEQ 203  
 QY 208 KMQPP 212  
 DB 204 PLEHP 208

RESULT 10  
 ID YN32\_MYCTU STANDARD; PRT; 291 AA.

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HYPOHETICAL 30.7 KDA PROTEIN RV2232/RV2233.  
 GN RV2232/RV2233 OR MT2292 OR MTCY427.13/MTCY427.14.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Filum: Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RA MEDLINE=98295987; PubMed=9634230; Garnier T., Churcher C., Harris D.,  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekle A.F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulten J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Swinn M.L., Hart D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains";  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CBBI/CBBZ/GPH/ITIE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL: 270692; CAA94666.1; ALT TERM.  
 DR EMBL: 270692; CAA94655.1; ALT INIT.  
 DR EMBL: A007074; AAK46576.1; -  
 DR TIGR: MT2292; -  
 DR TubercuList; RV2232; -  
 DR TubercuList; RV2233; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 291 AA; 30694 MW; 750F09FB154E6B5 CMC64;

Query Match 6.5%; Score 79; DB 1; Length 2  
 Best Local Similarity 25.8%; Pred. No. 3.3;  
 Matches 47; Conservative 16; Mismatches 65; Indels 51; Gaps 11;

QY 23 PAVPALNRTGPGSPSIQKTYDLTRYLEHQLSLAGTY---LNYLGLPNNQDP----- 74  
 DB 64 PPRAAARRASSGSPSPOL-VIPDLDTLTDSARGIVSFRRLNHTGAPVPRDLATHIV 122  
 QY 75 NPPRLCAETLP-----RATVLEWRSINDRLRTQNEAVSHLCTYLR- 118  
 DB 123 GPPM--HETLRAMGLGESEAEIAYRADYSARGW-----AMNSLPFCIPIIADLRT 173  
 QY 119 -GLNQOATAE---LRLSLAF-----CTSLGGLGSIAGVATLGYPLPQ--P 161  
 DB 174 AGVRLAVATSKAPPARLRLRHFQHEFVYAGASTDSGSKVDVLA--LVAQLHP 230  
 QY 162 LP 163  
 DB 231 LP 232

RESULT 11  
 ID BAT3\_HUMAN STANDARD; PRT; 1132 AA.

AC P46379;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE LARGE PROLINE-RICH PROTEIN BAT3 (HLA-B-ASSOCIATED TIVANSCN-1 3).  
 GN BAT3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CELL;  
 RA MEDLINE=90192810; PubMed=2156268;  
 RA Banerji J., Sands J., Strominger J.L., Spies T.;  
 RA "A gene pair from the human major histocompatibility complex encodes  
 RT large proline-rich proteins with multiple repeated motifs and a  
 RT single ubiquitin-like domain.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).  
 RN [2]  
 RP FUNCTION: UNKNOWN.

CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL: M33519; AAA35587.1; -  
 DR EMBL: M33521; AAA35588.1; -  
 DR EMBL: M33520; AAA35588.1; JOINED.  
 DR MIM: 142590; -



DR InterPro: IP000626; Ubiquitin.  
 DR Pfam: PF00240; ubiquitin; 1.  
 DR SMART: SM00213; UBO; 1.  
 DR PROSITE: PS00299; UBIQUITIN\_1; 1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 KW Repeat.  
 FT DOMAIN 17 77 UBIQUITIN-LIKE.  
 FT POLY-PRO.  
 FT DOMAIN 202 207 4 X 29 AA APPROXIMATE REPEATS.  
 FT REPEAT 242 270 1.  
 FT REPEAT 415 443 2.  
 FT REPEAT 574 602 3.  
 FT REPEAT 608 636 4.  
 FT DOMAIN 657 670 POLY-PRO.  
 SQ SEQUENCE 1132 AA; 119504 MW; E28CABA78C38DD18 CRC64;

Query Match 6.5%; Score 79; DB 1; Length 1132;  
 Best Local Similarity 26.1%; Pred. No. 18;  
 Matches 46; Conservative 17; Mismatches 67; Indels 46; Gaps 11;

OY 23 PAVNALNRGDPGPGPSIOKTYDLRLRYLEHQLSLAGTYLWYLGPPNEDPFRIGAE 82  
 DB 261 PA-PEYAPNHPSPAEYVEVLQELQR-LESRLQPLQRYEVLGAAP-TYNNHGRE 317  
 OY 83 TLPRATVLEWRSILNRLRLTON-YEAYSHLCLYLRGLNQAAT---AELRSIAHFC 137  
 DB 318 EDQR-LINL-----VGSRLRLGTMFVALSRLCNL-----ACPRPHLHYVRMSHYT 365  
 OY 138 TSLQGLGSIAGVWATLGYPL-----FQPLPGE-PAWAGPAHS 176  
 DB 366 TPM-----VLQQAAPLPIQINVGTVTMTGNGTRPPTPNAPPPGQASS 412

RESULT 12  
 ATKA\_MYCTU STANDARD; PRT; 571 AA.  
 ID ATKA\_MYCTU  
 AC P96371;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE POTASSIUM-TRANSPORTING ATPASE (EC 3.6.3.12) (POTASSIUM-TRANSPORTING ATPASE A CHAIN) (ATP PHOSPHOHYDROLASE [POTASSIUM-TRANSPORTING] A CHAIN) (POTASSIUM BINDING AND TRANSLOCATING SUBUNIT A).  
 DE KOPA OR RV1029 OR MT1058 OR MTCY10G2.20C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV.  
 RX MEDLINE=9825987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Stulson J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence".  
 RT Nature 393:537-544(1998).  
 RL Nature 393:537-544(1998).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COC 1551 / Oshkosh;  
 RA Fleischnann R.D., Allard D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolony J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D., Salzberg S.L., Salzberg A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishal W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains".  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ database.  
 CC -!- FUNCTION: ONE OF THE COMPONENTS OF THE HIGH-AFFINITY ATP-DRIVEN POTASSIUM TRANSPORT (OR KDP) SYSTEM, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF HYDROGEN AND POTASSIUM IONS (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + K(+)(OUT) = ADP + PHOSPHATE + K(+)(IN).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PRO-STATE).  
 CC -!- SIMILARITY: BELONGS TO THE KOPA FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by individuals for commercial entities requires a license agreement (see [http://www.ish.sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@sib.ch](http://www.ish.sib.ch/announce/or_send_an_email_to_license@sib.ch)).  
 CC EMBL: Z92539; CAB06850.1; -.  
 CC EMBL: AE006988; AAK45310.1; -.  
 DR EMBL: Z92539; CAB06850.1; -.  
 DR TIGR: MT1058; -.  
 DR Tuberculist; RV1029; -.  
 KW Hydrolyase; Transport; Potassium transport; Transmembrane; Complete proteome.  
 FT TRANSMEM 7 27 POTENTIAL.  
 FT TRANSMEM 66 86 POTENTIAL.  
 FT TRANSMEM 137 157 POTENTIAL.  
 FT TRANSMEM 188 208 POTENTIAL.  
 FT TRANSMEM 255 275 POTENTIAL.  
 FT TRANSMEM 286 306 POTENTIAL.  
 FT TRANSMEM 369 389 POTENTIAL.  
 FT TRANSMEM 390 410 POTENTIAL.  
 FT TRANSMEM 430 450 POTENTIAL.  
 FT TRANSMEM 497 517 POTENTIAL.  
 FT TRANSMEM 539 559 POTENTIAL.  
 SQ SEQUENCE 571 AA; 60163 MW; 4A3238D1C62A3687 CRC64;

Query Match 6.4%; Score 78.5; DB 1; Length 1132;  
 Best Local Similarity 22.9%; Pred. No. 8.7;  
 Matches 54; Conservative 31; Mismatches 76; Indels 75; Gaps 12;

OY 5 AGDSWGLACITCTVLMHPAYPAL-----NRTGDPGPGPSIOK 46  
 DB 2 SGTSMLOFRALIAVL--LTPALGGLYLAITYGDEAKKRGDVFGEIVNT-VCRVDEG 59  
 OY 47 --TRYLEHQLSLAGTYLNYL-----GPPNEDPFRIGAE 137  
 DB 60 SEQRSTVALSVLAFSVMSEFLLYGLARFGVLPENPTD--KPAVTDIYA 117  
 OY 94 --WRSILNRLRLT-----QNYEASHLCTL-----RGINQQA 131  
 DB 118 TNWQSYSGEPAMSHFTQMTGLAVQNFVSASAGCVIALIRGLARKRAS 177  
 OY 132 SLAHCTSLP-----QGLGISIAG-VMAITGLYPLPQPGTEPAMAI 176  
 DB 178 TYLRIMFLSFVALLVLSQGVIONLHGRTIVANTLEGAGQLLPG-----AVAS 226

RESULT 13  
 POLR\_TYMW STANDARD; PRT; 1844 AA.  
 ID POLR\_TYMW  
 AC P10358;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE RNA REPLICASE POLYPROTEIN (EC 2.7.7.48).  
 OS Turnip yellow mosaic virus.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage.  
 OX NCBI\_TaxID=12154;  
 RN [1]

```

RE SEQUENCE FROM N.A.
RX MEDLINE-88289359; PubMed-3399388;
RA Moch M.D., Boyer J.C., Haenni A.L.;
RT "Overlapping open reading frames revealed by complete nucleotide
RL sequencing of turnip yellow mosaic virus genomic RNA.";
RU Nucleic Acids Res. 16:6157-6173(1988).
CC CC CELLAEROUS: THE 206 KDA PROTEIN OF THE SINDBIS VIRUS.
CC (BY HOMOLOGU TO THE LONGER PROTEIN OF THE SINDBIS VIRUS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X07441; CAA30322.1; ALT_SEQ.
DR PIR: S01956; S01956.
DR MEROPS: C21.001; -.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF01443; Viral_helicase1.
KM Transferrase: RNA-directed RNA polymerase: Polyprotein; ATP-binding.
FT NP_BIND 976 983 ATP (BY SIMILARITY).
FT DOMAIN 1631 1663 POLYMERASE SITE (BY SIMILARITY).
SQ SEQUENCE 1844 AA; 206640 MW; A016D758C83D128C CRC64;

Query Match 6.4%; Score 78.5; DB 1; Length 1844;
Best Local Similarity 24.6%; Pred. No. 38;
Matches 57; Conservative 18; Mismatches 60; Indels 97; Gaps 13;

QY 7 DSMGMLACLTVMHLPANPALNRTGDPGPGSID-----KTYDLTR-YLEHOLSLAGT 60
DB 237 ESMG-----PVHSLIDRGTPRPDSQAPRTMTSDFRSYQGERLDVVS-- 282
QY 61 YLNLGPPNEPDNPRLGAETLPRAVNLVWRSLSLDRILTONYATYSHLLCYLRL 120
DB 283 -----FRIPD-----AIEPQATF---LQPLDRKVRPRAVNA---LFYTR-- 319
QY 121 NROATAELRSLAHFTSLQGLGSLAGVATLGYPLPPLPTEPAPGAPSAHSDFLQ 180
DB 320 -----AVRLKRS-----DPA-AVVRHMS---S 338
QY 181 KMDFFWLKELQTLWRSKXDFNRLKKMP-----PAASYLHLLEAH 223
DB 339 KPDDHWTSN-----AMDNLQTFALLNPLRPVNYHYLQSPLASLSTYLROH 386

RESULT 14
Y140_HUMAN STANDARD; PRT; 422 AA.
AC Q14153;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN KIAA0140.
GN KIAA0140.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE-96127530; PubMed-8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIA0121-KIAA0160) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:167-174(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D50930; BAA09489.1; -.
KM Hypothetical protein.
SQ SEQUENCE 422 AA; 45796 MW; 8F607F093408CEAC CRC64;

Query Match 6.4%; Score 78; DB 1; Length 422;
Best Local Similarity 21.6%; Pred. No. 6.6;
Matches 45; Conservative 21; Mismatches 74; Indels 68; Gaps 8;

QY 4 RAGDSMGMLACLTVMHLPANPALNRTGDPGPGSIDKTYDLTRYL-EHOLSLAGT 61
DB 206 QAGDWS-----PDLHPVG-----GRDLQSLSCSHQPS-----F 238
QY 62 LNYLGPPEPDNPRLGAETLPRAVNLVWRS-----LNR----- 100
DB 239 VEYCPSPANSSTPASTPELA-----RRSSGLSRSPQPCVLDKKGVKKRRPPEPQORP 293
QY 101 -----LRLTONYAYSHLLCYLGNLROAATLRLSLAHFTSLQGLGSLAGVATLGY 156
DB 294 SLDLAKMAQNCQTFSSLSCLSGATEDCGPQSPFAHVSNTRAWLALLSASNGRTPAGT 353
QY 157 PLPPLPTEPAPGAPSAHSDFLQKMD 184
DB 354 PVPEPLP-----PSFDHLVCOED 372

RESULT 15
RT65_MYXXA STANDARD; PRT; 427 AA.
AC P23071;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RNA-DIRECTED DNA POLYMERASE FROM RETRON MX65 (EC 2.7.7.7) (NEVERSE
DE TRANSCRIPTION) (MX65-RT).
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacterales;
OC Myxococcales; Cytophacterales; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90138993; PubMed-1689062;
RA Inouye S., Herzer P.J., Inouye M.;
RT "Two independent retrans with highly diverse reverse transcriptases
RT in Myxococcus xanthus.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:942-945(1990).
CC CC FUNCTION: PARTICIPATES IN THE SYNTHESIS OF A MSDNA (A BRANCHED RNA
CC LINKED BY A 2',5'-PHOSPHODIESTER LINKAGE TO A SINGLE-STRANDED DNA)
CC OF UNKNOWN FUNCTION. THE RETRON TRANSCRIPT SERVES AS PRIMER AND
CC TEMPLATE TO THE REACTION, AND CODES FOR THE RT.
CC CC CATALYTIC ACTIVITY: N-DEOXYNUCLEOSIDE TRIPHOSPHATE
CC N PYROPHOSPHATE + DNA(n).
CC CC MISCELLANEOUS: M.XANTHUS CONTAINS TWO INDEPENDENT AND UNLINKED
CC RETRONS: MX65 AND MX162.
CC CC MISCELLANEOUS: RETRONS MAY BE THE ANCESTORS OF RETROVIRUS.
CC CC -1- SIMILARITY: TO OTHER REVERSE TRANSCRIPTASES IN BACTERIA AND
CC RETROVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

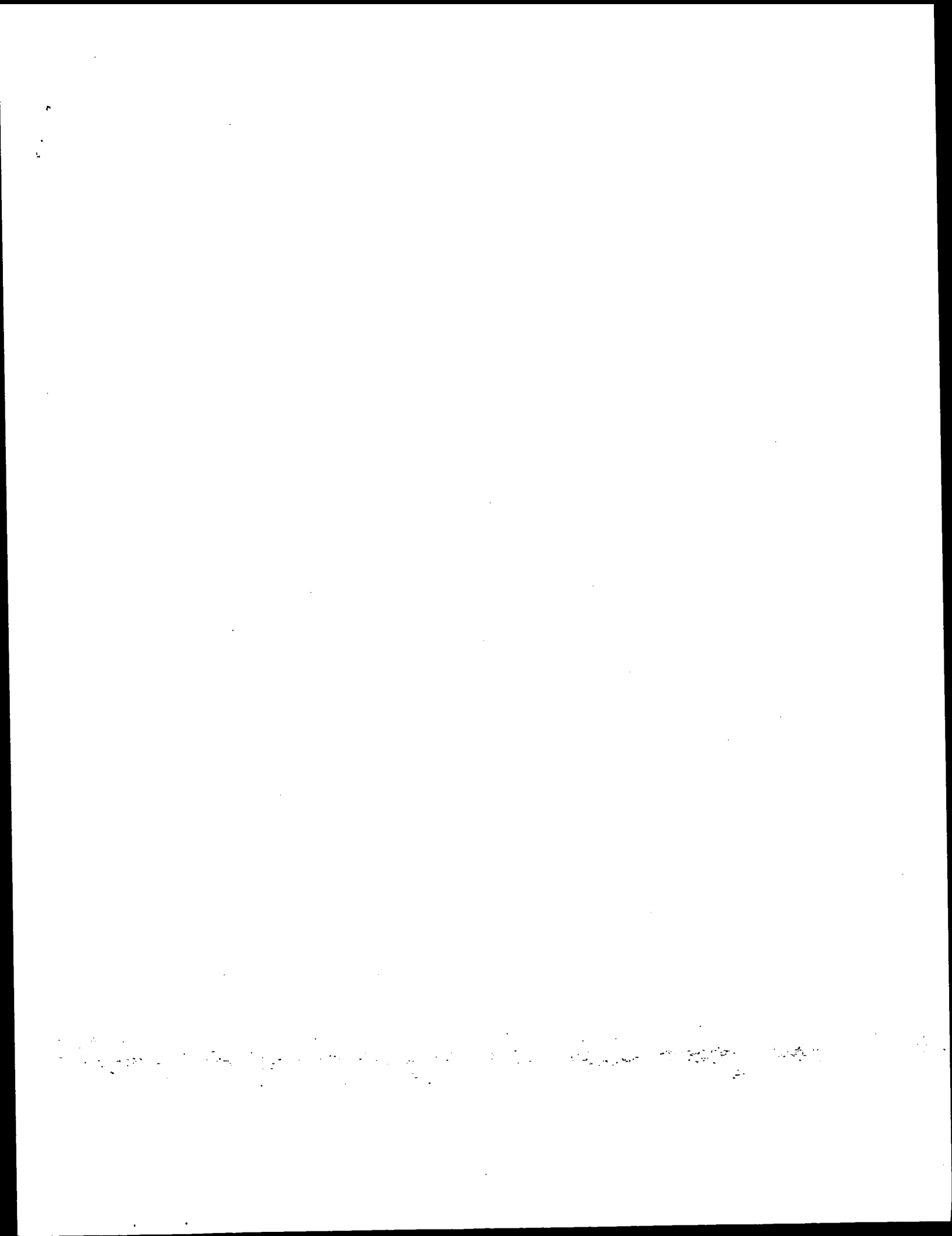
```

DR EMBL; M30609; AAA88323.1; .  
 DR PIR; A34864; RRYC65.  
 DR InterPro; IPR000123; RNA\_DNAPolym.  
 DR InterPro; IPR000477; RYTSe.  
 DR Pfam; PF00078; rvt; 1.  
 DR PRINTS; PR00866; RNADNAPOLMS.  
 KW Transferase; RNA-directed RNA polymerase; Transposable element.  
 SQ SEQUENCE 427 AA; 48024 MW; C96436BF86E4D0D7 CRC64;

Query Match 6.4%; Score 78; DB 1; Length 427;  
 Best Local Similarity 24.4%; Pred. No. 6.7;  
 Matches 38; Conservative 23; Mismatches 53; Indels 42; Gaps 6;

QY 18 VLMHLPAPALNRTGDPG-----PGPSIOKTYDLYRLE----- 51  
 Db 266 ILFHVPGPRVCYOGAPTSPALCNVALLRLDRRLAGLARRGYTYRYADDLTFSGDDVT 325  
 QY 52 --HQLSLAGTYLNYIGPPNPDPNPRL-GAETLPRAVNLVWRSNDRLRL----- 103  
 Db 326 ALERVRLAARYVOEGEFVNRKTRVQRGGAQRVTGVTNTTGLSREERPRRLAMLH 385  
 QY 104 ---TONYEAY-SH---LICYLRLGLNRQAATFELRR 131  
 Db 386 QEARSEDVEHRAHLDDLALAYMNLNPEQARLARR 421

Search completed: March 18, 2002, 09:40:14  
 Job time: 530 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:39:15 ; Search time 75.65 Seconds  
(without alignments)  
435,047 Million cell updates/sec

Title: us-09-931-704-5  
Perfect score: 1224  
Sequence: 1 MDLRAGDSKGMACICTYLVW.....KKKMPASVTLHLAHGF 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP:archaea:\*  
2: SP:bacteria:\*  
3: SP:fungi:\*  
4: SP:human:\*  
5: SP:invertebrate:\*  
6: SP:mammal:\*  
7: SP\_mhc:\*  
8: SP:organelle:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-rodent:\*  
12: SP-virus:\*  
13: SP-vertebrate:\*  
14: SP-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1224	100.0	225	11 09QZM3	Q9gzm3 mus musculu
2	1193	97.5	225	4 09UBP9	Q9ubp9 homo sapien
3	144.5	11.8	215	13 09PUJ2	Q9puj2 plethodon j
4	144.5	11.8	215	13 09PUJ1	Q9puj1 plethodon j
5	144.5	11.8	215	13 09PUJ0	Q9puj0 plethodon j
6	142.5	11.6	215	13 09PUJ9	Q9puj9 plethodon j
7	93.5	7.6	640	2 054153	054153 streptomyc
8	93	7.6	727	11 088841	088841 mus musculu
9	92	7.5	727	10 09MAU1	09mau1 arabidopsis
10	91.5	7.5	455	11 09CWV7	09cwv7 mus musculu
11	91	7.4	423	11 09JHE4	09jhe4 mus musculu
12	90	7.4	423	11 09JHE5	09jhe5 pseudomonas
13	89	7.3	423	11 09DBV6	09dbv6 mus musculu
14	89	7.3	733	2 091664	091664 pseudomonas
15	88	7.2	222	2 09L584	09l584 streptococ
16	87	7.1	1620	2 09Y4K6	09y4k6 homo sapien
17	86.5	7.1	1429	4 09HAV8	09hav8 homo sapien
18	86	7.0	559	4 099799	099799 homo sapien
19	86	7.0	619	4 090QJ2	090qj2 homo sapien

20	86	7.0	619	4 09UBU9	Q9ubu9 homo sapien
21	85.5	7.0	389	1 09YA71	Q9ya71 aeropyrum p
22	85.5	7.0	1154	11 0921R2	Q921r2 mus musculu
23	85	6.9	520	2 09RT14	Q9rt14 deinococcus
24	85	6.9	585	5 09VNR6	Q9vnr6 atrophilia
25	84	6.9	741	2 09H2M0	Q9h2m0 pseudomonas
26	83	6.8	272	2 086721	086721 streptomyc
27	83	6.8	396	2 09AF00	Q9af00 frankia sp.
28	83	6.8	955	11 088287	088287 mus musculu
29	83	6.8	1561	11 088286	088286 rattus norv
30	82.5	6.7	1098	11 09WTN8	Q9wtm8
31	82.5	6.7	5990	2 09RLP6	Q9rlp6 mycobacteri
32	81.5	6.7	209	2 09LS93	Q9ls93 streptococ
33	81.5	6.7	256	2 09LS90	Q9ls90 streptococ
34	81.5	6.7	294	11 09CP21	Q9cp21 mus musculu
35	81	6.6	227	2 09KGS0	Q9kgs0 streptococ
36	81	6.6	237	2 09LS92	Q9ls92 streptococ
37	81	6.6	256	2 09LS95	Q9ls95 streptococ
38	81	6.6	257	2 09LS94	Q9ls94 streptococ
39	81	6.6	395	2 09LAV9	Q9lav9 neurospora
40	81	6.6	1266	3 09ZCD2	Q9zcd2 streptomyc
41	80.5	6.6	373	2 09ZGC0	Q9zgc0 streptomyc
42	80.5	6.6	379	2 09ZGH4	Q9zgh4 streptomyc
43	80	6.5	294	4 09NXX5	Q9nxx5 homo sapien
44	80	6.5	876	4 09P210	Q9p210 homo sapien
45	80	6.5	1179	10 065507	065507 arabidopsis

## ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	225 AA.
09QZM3				
ID	09QZM3			
AC	09QZM3			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	NEUROTROPHIN-1/B-CELL STIMULATING FACTOR-3.			
GN	BSF3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus			
OX	NCBI_TaxID=10090;			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99432254; PubMed=10500198;			
RA	Sennaldi G., Varnum B.C., Sarmiento U., Starnes C., Liu			
RA	Guo J., Elliott G., McNinch J., Shaklee C.L., Freeman L.			
RA	Simonet W.S., Boone T., Chang M.-S.;			
RT	"Novel neurotrophin-1/B cell-stimulating factor-3: A cy			
RT	IL-6 family."			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:11458-11463(1999).			
DR	EMBL: AF176913; AF00993.1; -			
DR	MGD: MGI:1930088; BSf3.			
SQ	SEQUENCE 225 AA; 25261 MW; 68BJFEABVFLA950 CRG61.			

Query Match	100.0%;	Score 1224;	DB 11;	L=225;
Best Local Similarity	100.0%;	Pred. No. 1.1e-110;		
Matches 225; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	MDLRAGDSKGMACICTYLVWHPALNRTGDPGSPSTQKTYDLTK	1	QKSLACT 60
DB	1	MDLRAGDSKGMACICTYLVWHPALNRTGDPGSPSTQKTYDLTK	1	QKSLACT 60
QY	61	YLVYLGPPNEPDEPNRPGATLPRAVTVNEVMSLNDRLRLQNFVAVS	61	LCYIRGL 120
DB	61	YLVYLGPPNEPDEPNRPGATLPRAVTVNEVMSLNDRLRLQNFVAVS	61	LCYIRGL 120
QY	121	NROAAFAELRRLAHPCSTLQGLISAGVMTTCPLPPLPCTEPAVA	121	HSDFLO 180
DB	121	NROAAFAELRRLAHPCSTLQGLISAGVMTTCPLPPLPCTEPAVA	121	HSDFLO 180

QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 225  
 DB 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 225

RESULT 2  
 ID 09UBD9 PRELIMINARY: PRT; 225 AA.

AC 09UBD9; 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DE NEUROTOPHIN-1/B-CELL STIMULATING FACTOR-3.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
 RX MEDLINE-99432254; PubMed-10500198;  
 RA Senaldi G., Varnum B.C., Sarmiento U., Starnes C., Lille J., Scully S.,  
 RA Guo J., Elliott G., McNinch J., Shaklee C.L., Freeman D., Manu F.,  
 RA Simonet W.S., Boone T., Chang M.-S.;  
 RT "Novel neurotrophin-1/B cell-stimulating factor-3: A cytokine of the  
 RT IL-6 family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11458-11463(1999).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE-99382254; PubMed-10448081;  
 RA Shi Y., Wang W., Yourey P.A., Gohari S., Zukauskas D., Zhang J.,  
 RA Ruben S., Alderson R.F.;  
 RT "Computational EST database analysis identifies a novel member of the  
 RT neurotrophic cytokine family.";  
 RL Biochem. Biophys. Res. Commun. 262:132-138(1999).  
 DR EMBL: AF172854; AAS4284.1; -;  
 DR EMBL: AF176911; AAF00991.1; -;  
 SQ SEQUENCE 225 AA; 25176 MW; E2DD4B628083B55 CRC64;

Query Match  
 Best Local Similarity 97.5%; Score 1193; DB 4; Length 225;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGLACLTCTVHLPAVPAVLRNRTGPGPSTQKTYDRLTRYEHQLRSLAGT 60  
 DB 1 MDLRAGDSWGLACLTCTVHLPAVPAVLRNRTGPGPSTQKTYDRLTRYEHQLRSLAGT 60  
 QY 61 YLNYLGPPEPPENPPRRLGAEPLPRATVNLVWRSLSNDRLLRLTONYEYSHLLCYLRGL 120  
 DB 61 YLNYLGPPEPPENPPRRLGAEPLPRATVNLVWRSLSNDRLLRLTONYEYSHLLCYLRGL 120  
 QY 121 NRGAAATAEELRSLAHFCTSLQGLGSIAGVMATIGYPLPQPLGTPPAARPAASDFLQ 180  
 DB 121 NRGAAATAEELRSLAHFCTSLQGLGSIAGVMATIGYPLPQPLGTPPAARPAASDFLQ 180  
 QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 225  
 DB 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 225

RESULT 3  
 ID 09PUJ2 PRELIMINARY: PRT; 215 AA.  
 AC 09PUJ2; 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DE RECEPTIVITY FACTOR ISOFORM 1 PRECURSOR.  
 GN PRe.  
 OS Plethodon jordanii (Salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;  
 OC Plethodon.  
 NCBI\_TaxID=8336;

RP SEQUENCE FROM N.A.  
 RX MEDLINE-99420364; PubMed-10489368;  
 RA Rollmann S.M., Houck L.D., Feldhoff R.C.;  
 RT "Proteinaceous pheromone affecting female receptivity in the terrestrial  
 RT salamander.";  
 RL Science 285:1907-1909(1999).  
 DR EMBL: AF181480; AAF01025.1; -;  
 KW Signal.  
 FT SIGNAL. 1 23 POTENTIAL.  
 SQ SEQUENCE 215 AA; 24138 MW; B1906B866335738 CRC64;

Query Match  
 Best Local Similarity 11.8%; Score 144.5; DB 13; Length 215;  
 Matches 41; Conservative 31; Mismatches 81; Indels 7; Gaps 3;

QY 56 SLAGTYNYLGPPEPPENPPRRLGAEPLPRATVNLVWRSLSNDRLLRLTONYEYSHLLC 115  
 DB 56 SLAGTYNYLGPPEPPENPPRRLGAEPLPRATVNLVWRSLSNDRLLRLTONYEYSHLLC 115  
 QY 116 YLR-GLNRO----ATAELRSLAHFCTSLQGLGSIAGVMATIGYPLPQPLGTPPA 170  
 DB 116 YLR-GLNRO----ATAELRSLAHFCTSLQGLGSIAGVMATIGYPLPQPLGTPPA 170  
 QY 171 PGPAHSDFLQKMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 210  
 DB 171 PGPAHSDFLQKMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 210  
 QY 173 PGPAHSDFLQKMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 212  
 DB 173 PGPAHSDFLQKMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 212

RESULT 4  
 ID 09PUJ1 PRELIMINARY: PRT; 215 AA.  
 AC 09PUJ1; 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DE RECEPTIVITY FACTOR ISOFORM 2 PRECURSOR.  
 GN PRe.  
 OS Plethodon jordanii (Salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_TaxID=8336;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99420364; PubMed-10489368;  
 RA Rollmann S.M., Houck L.D., Feldhoff R.C.;  
 RT "Proteinaceous pheromone affecting female receptivity in the terrestrial  
 RT salamander.";  
 RL Science 285:1907-1909(1999).  
 DR EMBL: AF181481; AAF01026.1; -;  
 KW Signal.  
 FT SIGNAL. 1 23 POTENTIAL.  
 SQ SEQUENCE 215 AA; 24080 MW; B341B8B7B4E28438 CRC64;

Query Match  
 Best Local Similarity 11.8%; Score 144.5; DB 13; Length 215;  
 Matches 41; Conservative 31; Mismatches 81; Indels 7; Gaps 3;

QY 56 SLAGTYNYLGPPEPPENPPRRLGAEPLPRATVNLVWRSLSNDRLLRLTONYEYSHLLC 115  
 DB 56 SLAGTYNYLGPPEPPENPPRRLGAEPLPRATVNLVWRSLSNDRLLRLTONYEYSHLLC 115  
 QY 116 YLR-GLNRO----ATAELRSLAHFCTSLQGLGSIAGVMATIGYPLPQPLGTPPA 170  
 DB 116 YLR-GLNRO----ATAELRSLAHFCTSLQGLGSIAGVMATIGYPLPQPLGTPPA 170  
 QY 171 PGPAHSDFLQKMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 210  
 DB 171 PGPAHSDFLQKMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLLEAHGF 210

Db 173 PFSSAIFRKKKRGVCKEYKERVLTTRDFELAKKYQ 212

## RESULT 5

09PU10 PRELIMINARY; PRT; 215 AA.

AC 09PU10: 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE RECEPTIVITY FACTOR ISOFORM 3 PRECURSOR.  
 GN PRF.  
 OS Plethodon jordani (Salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;  
 OC Plethodon.  
 OC NCBI\_TaxID=8336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99420364; PubMed=10489368;  
 RA Kollmann S.M., Houck L.D., Feldhoff R.C.;  
 RT "Proteinaceous pheromone affecting female receptivity in a terrestrial salamander."  
 RL Science 285:1907-1909(1999).  
 DR EMBL: AF181482; AAF01027.1; -.  
 KW Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 SQ SEQUENCE 215 AA; 24024 MW; 6EB6E4AD93A1343 CRC64;

Query Match 11.8%; Score 144.5; DB 13; Length 215;  
 Best Local Similarity 26.5%; Pred. No. 3.4e-06;  
 Matches 44; Conservative 34; Mismatches 73; Indels 15; Gaps 4;

QY 56 SLAGTYLNTLGPFPNPPRLGAEPLPRATVNLVWRSINDRLRLTONYEAYSHLLC 115  
 DB 55 SLPLTYLSFGAPLSDPDYQLPHIKVYNLPTAMDDYDFIRQTDERRKNNLYFSAYIE 114  
 QY 116 YLR-----GLN--RQAATALRLSLAHCSTSLGSLGAGVATLGYPLPLPGTE 166  
 DB 115 FLKEAMTEQEDLNPAELALAKFEAMANSNT---LISKISIDIMQMSVITLP--K 168  
 QY 167 PAMAPGPAHSDFLQKMDDFMLKELQTLWLRSAKDNRLKKMQP 212  
 DB 169 PLVPPKGSAYFSKRLGVCYCKERYVFLTRDFMLLAKEYQ 214

## RESULT 6

09PU10 PRELIMINARY; PRT; 215 AA.

AC 09PU10: 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE RECEPTIVITY FACTOR ISOFORM 4 PRECURSOR.  
 GN PRF.  
 OS Plethodon jordani (Salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;  
 OC Plethodon.  
 OC NCBI\_TaxID=8336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99420364; PubMed=10489368;  
 RA Kollmann S.M., Houck L.D., Feldhoff R.C.;  
 RT "Proteinaceous pheromone affecting female receptivity in a terrestrial salamander."  
 RL Science 285:1907-1909(1999).  
 DR EMBL: AF181483; AAF01028.1; -.  
 KW Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 SQ SEQUENCE 215 AA; 24054 MW; AAA412135FRA4E7F CRC64;

Query Match 11.6%; Score 142.5; DB 13; Length 215;  
 Best Local Similarity 25.3%; Pred. No. 5.3e-06;  
 Matches 41; Conservative 33; Mismatches 81; Indels 7; Gaps 3;

QY 56 SLAGTYLNTLGPFPNPPRLGAEPLPRATVNLVWRSINDRLRLTONYEAYSHLLC 115  
 DB 55 SLPLTYLSFGAPLSDPDYQLPHIKVYNLPTAMDDYDFIRQTDERRKNNLYFSAYIE 114  
 QY 116 YLR-GLNQ----ATAELRLSLAHCSTSLGSLGAGVATLGYPLPLPGTEPAMA 170  
 DB 115 FLKEAMTEQEDLNPAELALAKFEAMANSNTLISKISIDIMQMSVITLP--KPLV 172  
 QY 171 PPAHSDFLQKMDDFMLKELQTLWLRSAKDNRLKKMQP 212  
 DB 173 PFKGSAYFSKRLGVCYCKERYVFLTRDFMLLAKEYQ 214

## RESULT 7

054153 PRELIMINARY; PRT; 640 AA.

AC 054153: 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE TRANSFERASE.  
 GN SCF7.10.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OC NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Parkhill J., Barrell B.G., Rastam M.A.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A.;  
 RT "A set of ordered cosmids and a detailed genetic map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL021409; CAAL6181.1; -.  
 DR HSP: P12998; JBSO.  
 DR InterPro: IPR001917; AminoTransf.2.  
 DR InterPro: IPR003880; Phosphopant\_attach.  
 DR Pfam: PF00222; aminotran\_2; 1.  
 DR PROSITE: PS50075; ACP\_DOMAIN; 1.  
 KW Phosphopantetheine; Transferase.  
 SQ SEQUENCE 640 AA; 68061 MW; E619FB5E55A613F CRC64;

Query Match 7.6%; Score 93.5; DB 2; Length 140;  
 Best Local Similarity 28.3%; Pred. No. 1.1; Indels 57; Gaps 9;  
 Matches 52; Conservative 12; Mismatches 63;

QY 22 LPVAPALNRTGDPGPGSIQKTYDRLTYLEHOL---SLAGTYLNTLGPFPNPPR 78  
 DB 100 LPVAPR-EGTGDVTEVVAAMRQYORHGLADLEG-----E 140  
 QY 79 LGAETPLPRATVNLVWRSINDRLRLTONYEAYSHLLCYLRGNKRAAT 138  
 DB 141 LGVDSVVLTSVVAEA---TERLGLT-----GAAPDA 179

QY 139 SLOGGLGSIAGVATLGLYPLPPLGTEPAPAPCPAHSDFLQKMDDFWLKELQWTLMRS 198  
 DB 180 ALRCL-----VAAAPGTAPEAPAPGAA-APAPGRSGNAP-----APGADWDHRS 225  
 QY 199 AKDF 202  
 DB 226 MKDF 229

RESULT 8  
 088841  
 ID 088841 PRELIMINARY; PRT: 727 AA.  
 AC 088841  
 DT 01-NOV-1998 (Tremblrel, 08, Created)  
 DT 01-NOV-1998 (Tremblrel, 08, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel, 17, Last annotation update)  
 DE FACIOGENITAL DYSPLASIA PROTEIN 2.  
 GN FGD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pasteris N.G., Gorski J.L.;  
 RT "Isolation, characterization and mapping of Fgd2, a homologue of the  
 RT facio-genital dysplasia (FGD1; Aarskog syndrome) gene."  
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF017368; AAC35430.1;  
 DR MGI; MGI:1347084; Fgd2.  
 DR InterPro: IPR001849; PH.  
 DR InterPro: IPR000219; RhogEF.  
 DR InterPro: IPR000306; Znf\_FIVE.  
 DR Pfam; PF01363; FYVE; 1.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF00621; RhogEF; 1.  
 DR SMART; SM00064; FYVE; 1.  
 DR SMART; SM00233; PH; 1.  
 DR SMART; SM00325; RhogEF; 1.  
 DR PROSITE; PSS0003; PH\_DOMAIN; 1.  
 DR PROSITE; PSS0003; PH\_DOMAIN; 1.  
 SO SEQUENCE 727 AA; 82099 MW; 5CDDFBFB905C5FC5 CRC64;

Query Match 7.6%; Score 93; DB 11; Length 727;  
 Best Local Similarity 24.8%; Pred. No. 1.4;  
 Matches 61; Conservative 32; Mismatches 93; Indels 60; Gaps 11;

QY 21 HLPVPAALNRTGDPGPGPSIOKTYDRLTRYLEHQL-----RSLAGT 60  
 DB 13 NLVAVFENNR--PGAPGSHSLSDPHIPEHQLSLSPFEWAPPVYKALKSEFRPVSR 70  
 QY 61 YLNTLGGPPN-----EPDFNPRLGAEFLPRATVNLFWRSINRLRLTONYEAYSH 112  
 DB 71 YLSLTKKLKSSGAMRSCQGVSP--GPE-----TQEPKRVNRLLETEGAYVARLH 122  
 QY 113 LL--CYRLGNLRQAATRLRLSLAHCTSLQGLGSIAGVMA--TLGVPPLPPLPGLTEPAP 169  
 DB 123 LLQVFFQELLRLRAG-----RSKAFPEDEVYKLIIFSNISSIRPHAQFLPE--LQRRVDM 176  
 QY 170 APGPAHSDFLQKMDDFWLKELQWTLMRSADDFNRLLKKKMOOPASVY 217  
 DB 177 AATPRIGDVYQKLAFLKMTSEYKFNERRAELLATWMDKS--QPFQEVVYTRIQSEASS 235  
 QY 218 LHLFAH 223  
 DB 236 LTLQHH 241

RESULT 9  
 09MAU1 PRELIMINARY; PRT: 332 AA.  
 AC 09MAU1  
 QY 09MAU1

DT 01-OCT-2000 (Tremblrel, 15, Created)  
 DT 01-OCT-2000 (Tremblrel, 15, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel, 17, Last annotation update)  
 DE F13M7.8 PROTEIN.  
 GN F13M7.8.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Eudicotyledons;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-CV. COLUMBIA;  
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Kwan A., Liu, S.,  
 RA Li J., Kremenetskaia I., Lueros J., Araujo R., Au M., Bruch V.,  
 RA Buehler E., Conway A., Dewar K., Feng J., Kim C., Kurtz J., Li Y.,  
 RA Palm C., Shinn P., Sun H., Davis R., Ecker J., Federspiel J.,  
 RA Theologis A.;  
 RT "The sequence of BAC F13M7 from Arabidopsis thaliana chromosome 1."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ database;  
 DR EMBL; AC004809; AAF40444.1;  
 DR InterPro: IPR002965; P-rich\_extensin.  
 DR PRINTS; PR01217; PRICHEXTENSIN.  
 SO SEQUENCE 332 AA; 36793 MW; 1BB6687141A070F4 CRC64;

Query Match 7.5%; Score 92; DB 10; Length 332;  
 Best Local Similarity 25.9%; Pred. No. 0.7;  
 Matches 57; Conservative 24; Mismatches 73; Indels 66; Gaps 12;

QY 23 PAVPALNRTGDPGPGPSIOKTYDRLTRYLEHQLSLAGTYLNT----- 64  
 DB 113 PSYTAGNLGYP--PRSF--TYDPGYEQRQWESLLQDFIRERNQIRPLPLGLGSPVG 169  
 QY 65 LGPPFNPDPNPRLGAEFLPRATVNLFWRSINRLRLTONYEAYSH--LGLNRQA 124  
 DB 170 LGPIRASPGFLQPRVA---PEPTSLDTSRNRKRSK-----DGLAVLVK--RKV 215  
 QY 125 ATAELRSP-----AHCTSLQGLGSIAGVATLGLYPLPPLP--LTPAPMA 170  
 DB 216 RITESSSLYSIGRSLKNGAHV-----GLQPRSGIKM---PLPKPLVGLVLEISVP 266  
 QY 171 PGPASDFLQKMDDFWLKELQWTLMRSADDFNRLLKKKMO 210  
 DB 267 DDPDESADDEKEDBEAVKOL-----SEKDL--LKRRIE 298

RESULT 10  
 09CWV7 PRELIMINARY; PRT: 455 AA.  
 AC 09CWV7  
 DT 01-JUN-2001 (Tremblrel, 17, Created)  
 DT 01-JUN-2001 (Tremblrel, 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel, 17, Last annotation update)  
 DE 2410003H12RIK PROTEIN.  
 GN 2410003H12RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





QY 142 GLLGAGVATLGGYPLPGLPTGPAPWAGPASHDFLOKMDPWLKELQTMWRSKAD 201  
 DB 185 DLLGACGMEQELGGLVLAQ---GVES--AAGKANATYLEAANFYSSR-----FDEAEQG 234  
 QY 202 ENRLKKMKOPPAASVTLHLER 222  
 DB 235 FKALQDVSOFPWLKETALYLQA 255

RESULT 13  
 Q9DBV6 PRELIMINARY; PRT; 423 AA.  
 AC 09DBV6;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
 DE CEREBROSIDE (3'-PHOSPHOADENYLISULFATE:GALACTOSYLKERAMIDE 3')  
 DE SULFOTRANSFERASE.  
 GN CST.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinchik S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,  
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 403:685-690(2001).  
 DR EMBL; AK007645; BAB25160.1;  
 DR MGI; MGI:1858277; Cst.  
 SO SEQUENCE 423 AA; 48824 MW; 045303CFECCE150A3 CRC64;

Query Match 7.3%; Score 89; DB 11; Length 423;  
 Best Local Similarity 23.1%; Pred. No. 1.8;  
 Matches 55; Conservative 21; Mismatches 66; Indels 96; Gaps 11;

QY 44 YDLFVYLEHQRSLAGTYLVN---LGP-----FNEP----- 72  
 DB 204 YDPSSYNAHYRNLLFPLGDDSLDPASPVQHLLEVERRHVLVLQYFDESLVLLQ 263  
 QY 73 -----DENPRLGAETLPRAVNLVWRSINDRLITQNEAVSHL 113  
 DB 264 ELLCWDLDKLVLYFKLNARSDPVRPLRGSELVLRAT-----ANLLD---VRLYRHNASFWR 317  
 QY 114 LCYLRGLNROA-ATAELRIS---LAHCTSLQGLLSIA-----GVMAVLGYP 157  
 DB 318 KVEAFGRNAREVAELRQANENHRHICIDCGAVGAELADSDAMQWOPGLGKISILGYN 377  
 QY 158 LPQPLPSTEPAMAPGPAHSDFLOKMDPWLKELQTMWRSKAD 204  
 DB 378 LKRSI-----GPHQEDLCROM-----LTPREIQYLSDLGAGNLTWKMLKPLRDFUR 422

RESULT 14  
 Q91664 PRELIMINARY; PRT; 733 AA.  
 AC Q91664;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)  
 DE HYPOTHETICAL PROTEIN PA0454.  
 GN PA0454.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Burkholder P.,  
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik I.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004483; AAG03843.1;  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 733 AA; 81538 MW; 580F9DBCB3909DB CRC6;

Query Match 7.3%; Score 89; DB 2; Length 423;  
 Best Local Similarity 25.8%; Pred. No. 3.6;  
 Matches 57; Conservative 26; Mismatches 64; Indels 96; Gaps 13;

QY 28 LNRTGDPGPGPSIQTYDLRYLE-----HQLRSLAGTYLVLYLGPVPRD---FNPP 77  
 DB 230 LNRIGHGHPKV-----SRYLKYFLAQDYHERASSHYPRNRLAENI---DVLFRQ 283  
 QY 78 RL-----GATLPRAVNLVWRSINDRLITQNEAVSHLQYLNKINQDAVLELR 131  
 DB 284 RLWQGGKACALARA-----ITLRQFPD-YA-----DRELALDLQA 320  
 QY 132 SLAHF-----CTSLQGL-----LGSINGVATLGGYPLPGLPTGPAPWAGPASHDFLOKMD 184  
 DB 321 SLEHLRQSNPAWKGGLRSLGALAAANTITDRKL-----AGASNPATADEOD 368  
 QY 185 FWLKLQTMWRSKAD-FNRLKKMKOPPAASVTLHLER 224  
 DB 369 SALDRSP-----RSLKDAFERLRQQLTP-----TSLFRHG 400

RESULT 15  
 Q9L584 PRELIMINARY; PRT; 222 AA.  
 AC Q9L584;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE PSPA (FRAGMENT).  
 GN PSPA.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=43;  
 RA Beall B.W.;  
 RT "PSPA sequence types from multiresistant pneumococci."  
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF255543; AAF68096.1;

Mon Mar 18 11:23:42 2002

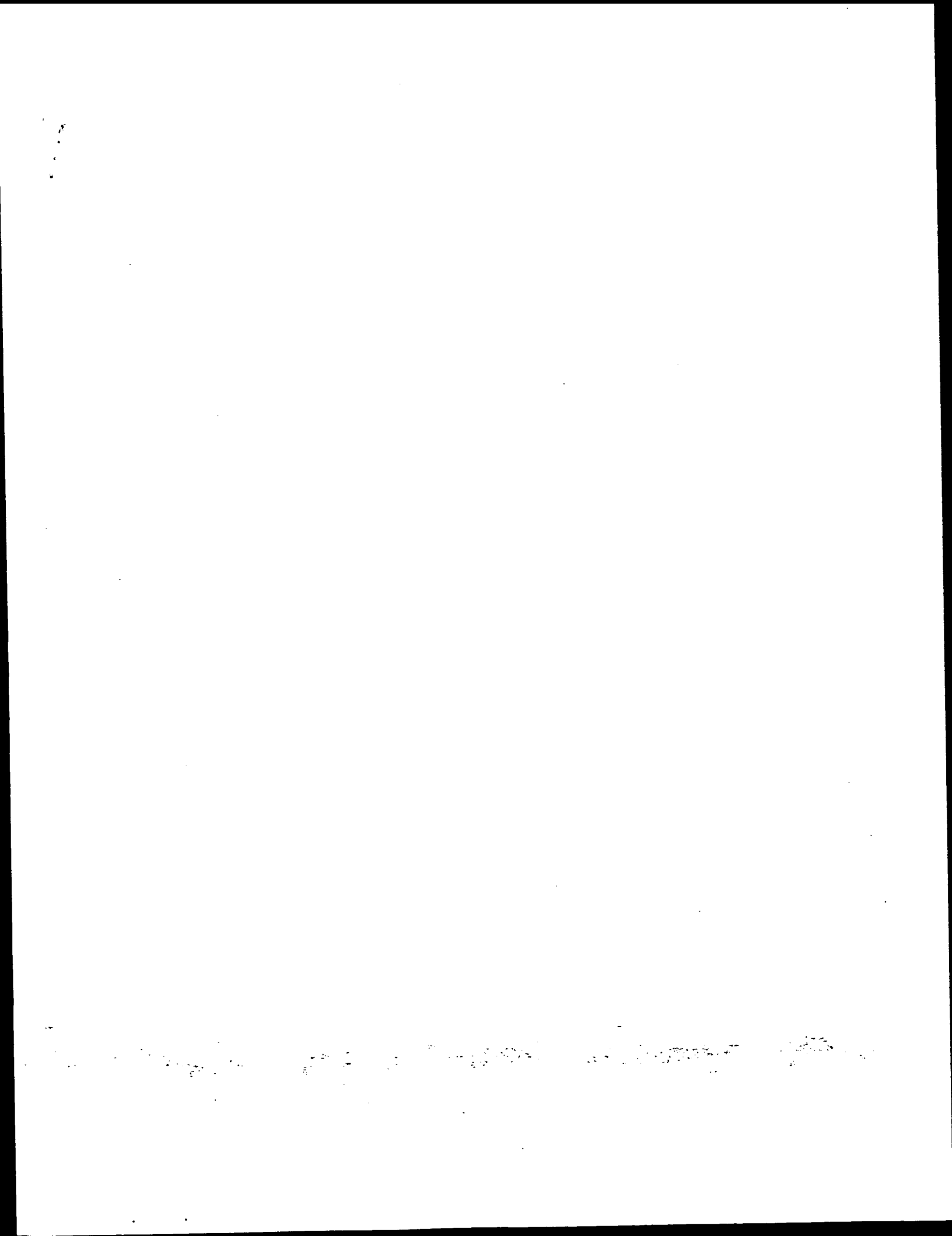
us-09-931-704-5.rspt

Page 7

```
DR      InterPro: IPR002965; P_rich_extensn.  
DR      PRINTS; PRO1217; PRICHEXTENSM.  
FT      NON_TER      1      1  
FT      NON_TER      222    222  
SQ      SEQUENCE      222 AA; 23952 MW; 48243
```

Query Match	7.28;	Score 88;	DB 2;	Length 222;
Best Local Similarity	27.98;	Pred. No. 1;		
Matches 39;	Conservative 16;	Mismatches 53;	Indels 32;	Gaps 5

```
OY 97 LNDRLRLQÑNEA-----YSHLCLCYGLNQAATLRLRSIAHCTSLQGLLSIAGVM 154
    ||::||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   65 LNKKVESLNQVADLEKEISINLETLLIGADEDEDTALQNKLATTAKALETEKQEDNAL 124
OY 152 ATLG-----YPLQLPLGTE-PAWAGPAHSDFLQMDPFWLKELQTW 194
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 NELPDDGEETPAAPDQEQPAAPAAPKPEQAPAPKPEKSADQQAEDY----- 175
OY 195 LMRSKADFNNLKKKQPAA 214
    ||:::|||::||| |
Db 176 ARSSEENRILTQG-QEPKA 194
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:30:21 ; Search time 55.47 Seconds

(without alignments)  
300.460 Million cell updates/sec

Title: US-09-931-704-5

Perfect score: 1224

Sequence: 1 MDLRAGDSWGMALACTVLW.....KKKMPAPASVTLHEAHGF 225

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_1101:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1224	100.0	225	19	AAW29716
2	1224	100.0	225	19	AAW56142
3	1224	100.0	225	21	AAW87814
4	1193	97.5	225	19	AAW29715
5	1193	97.5	225	19	AAW56141
6	1193	97.5	225	20	AAW94466
7	1193	97.5	225	21	AAW87813
8	1193	97.5	225	22	AAW56143
9	1193	97.5	225	22	AAW56143
10	1181	96.5	223	22	AAW56143
11	1167	95.3	215	21	AAW19587

12	1136	92.8	215	21	AAW19586
13	852	69.6	164	22	AAW20115
14	852	69.6	164	22	AAW34012
15	156.5	12.8	208	20	AAW09197
16	154	12.6	208	20	AAW09196
17	110.5	9.0	203	16	AAW39965
18	110.5	9.0	203	17	AAW88204
19	110.5	9.0	203	18	AAW29237
20	92.5	7.6	243	22	AAW20277
21	92	7.5	332	21	AAW22132
22	92	7.5	332	21	AAW22132
23	91.5	7.5	201	16	AAW39967
24	91.5	7.5	201	18	AAW29238
25	91.5	7.5	201	20	AAW06490
26	91.5	7.5	201	21	AAW27662
27	91.5	7.5	201	21	AAW13004
28	91.5	7.5	201	21	AAW3697
29	91.5	7.5	201	21	AAW87818
30	91.5	7.5	201	22	AAW50994
31	90	7.4	195	14	AAW34432
32	90	7.4	195	14	AAW3337
33	87	7.1	242	22	AAW20275
34	85.5	6.7	1473	22	AAW06758
35	82.5	6.7	241	16	AAW6166
36	81.5	6.7	348	22	AAW4996
37	81.5	6.7	352	14	AAW41000
38	81.5	6.6	134	18	AAW14585
39	81	6.6	379	21	AAW77185
40	80.5	6.6	620	21	AAW77208
41	80.5	6.6	379	21	AAW77179
42	80.5	6.6	3782	21	AAW77179
43	79.5	6.5	305	22	AAW95861
44	79.5	6.5	300	22	AAW03542
45	79.5	6.5	631	21	AAW43285

#### ALIGNMENTS

RESULT 1	
ID	AAW29716 standard; Protein; 225 AA.
XX	
AC	AAW29716;
XX	
DT	09-NOV-1998 (first entry)
XX	
DE	Mouse neurotrophic factor NNT-1.
XX	
KW	NNT-1; neurotrophic factor; mouse; antiinflammatory; adjuvant;
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW	amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW	peripheral neuropathy; dystrophy; neural retina degeneration;
KW	common variable immunodeficiency; CVID; selective iga deficiency;
KW	hypogammaglobulinemia; X-linked agammaglobulinemia; autisepic;
XX	
OS	Mus sp.
XX	
XX	
FT	Key
FT	Peptide
FT	Protein
XX	
XX	WO9833922-A1.
XX	
XX	06-AUG-1998.
XX	
XX	02-FEB-1998;
XX	
XX	02-FEB-1998;
XX	
XX	30-JAN-1998;
XX	
XX	03-FEB-1997;
PR	

Human interleukin-  
Peptide #6549 enco  
Peptide #8049 enco  
Human DNAX interle  
Human DNAX interle  
Mouse cardiac hype  
Human cardiac hype  
Human interleukin  
Human interleukin  
Arabidopsis thalia  
Human cardiac hype  
Human cardiac hype  
Human tumour-assoc  
Human protein PRO8  
Human cardiac hype  
Amino acid sequenc  
Human PRO82 prote  
Sequence of growth  
Chicken ciliary ne  
Human interleukin  
Polyhydroxyalkanoa  
Human G-protein co  
mml5. Mus muscula  
Human novel trans  
Human brain cDNA c  
St. leptococcus pneu  
St. venezuelae deso  
S. venezuelae deso  
S. venezuelae deso  
Human protein sequ  
Human protein kina  
Human ORFX ORF3049

Location/Qualifiers  
1..27  
/Label= Sig\_peptide  
28..225  
/Label= Mat\_protein

XX (AMGE-) AMGEN INC.

PI Chang M, Elliot GS, Sarmiento U, Senaldi G;

XX WPI; 1998-437475/37.

DR N-PSDB; AAV47512.

XX Newly isolated nucleic acid encoding human or murine neurotrophic  
PT factor NNT-1 - useful for treatment of neurological and  
PT immunological diseases or inflammation, also as vaccine adjuvant

PS Claim 13; Fig 5; 120pp; English.

XX This is the amino acid sequence of a murine neurotrophic factor,  
CC designated NNT-1, that is a growth factor for neurons and for B or  
CC T cells. It was deduced from isolated NNT-1 cDNA (see AAV47512).  
CC Human NNT-1 (see AAM29715) is also provided. Vectors and host cells  
CC for use in the production of human murine recombinant NNT-1  
CC polypeptides. These are used to treat: (1) neurological or  
CC immunological diseases, specifically Alzheimer's, Parkinson's  
CC or Huntington's diseases, amyotrophic lateral sclerosis,  
CC Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy and  
CC degeneration of the neural retina, or conditions characterised by T  
CC or B cell defects, e.g. common variable immunodeficiency (CVID),  
CC selective IgA deficiency, hypogammaglobulinaemia and X-linked  
CC agammaglobulinaemia (claimed), but many others disclosed; and (11)  
CC inflammation. NNT-1 is also able to boost immunoreactivity and  
CC antibody production following vaccination, and, since it inhibits  
CC tumour necrosis factor production, it may also be useful for  
CC treating sepsis. In addition, cells that have been engineered to  
CC express NNT-1 can be implanted, or nucleic acids are delivered in  
CC gene therapy vectors.

XX Sequence 225 AA:

Query Match 100.0%; Score 1224; DB 19; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.6e-121;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMCLACTCTVLMHLPAPALNRTGDPGPGPSIQTKYDITRYLEHOLRLSLAGT 60  
DB 1 mdlragsdswgmclactctvlmhpavpalnrtgdpgpsiqtkydltrylehqlrslagtl 60  
QY 61 YLNYLGPPEPNEPDPNPRIGAEETLPRAVNLKRWRSINDRLRLTONYEAVSHILCTLRGL 120  
DB 61 ylnylgppnepdpnprigaeetlpratvnlkwrslndrlrltonyeavshilcylrql 120  
QY 121 NRQATRAELRSLAHFCTSLQGLIGSTAGYMATLGYPLDPLPTEPMAWPGAHSDFLQ 180  
DB 121 nrqaataelrslahfctslqgligstagymatlgyplpplptepmawpgahsdflq 180  
QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHEAHGF 225  
DB 181 kmddfwlkeltlwlrskdfnrlkkkmpaasvvlheahgf 225

RESULT 2

AAW56142  
ID AAW56142 standard; Protein: 225 AA.

XX AAW56142;

XX 13-JUL-1998 (first entry)

XX Amino acid sequence of murine neurotrophic factor NNT-1.

XX Mouse; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;  
KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;  
KW amyotrophic lateral sclerosis; ALS; Alzheimer's disease; stroke.

XX Mus sp.

XX Key Location/Qualifiers  
FH Peptide 1..27  
FT Peptide /note= "signal peptide"  
FT Protein 28..225  
FT /note= "mature peptide"

XX US5741772-A.

XX 21-APR-1998.

XX 03-FEB-1997; 97US-0792019.

XX 03-FEB-1997; 97US-0792019.

XX (AMGE-) AMGEN INC.

XX Chang M;

XX WPI; 1998-260526/23.

XX N-PSDB; AAV22654.

PT Neurotrophic factor NNT-1 polypeptide and related nucleic acids  
PT useful for stimulating growth of motor and sympathetic neurons  
PS Claim 2; Fig 5; 41pp; English.

XX The present sequence represents a murine neurotrophic factor designated  
CC NNT-1, which is capable of stimulating growth of motor or sympathetic  
CC neurons. The NNT-1 protein is useful in the treatment of neurological  
CC diseases characterised by the degeneration and death of particular  
CC classes of neurons. These diseases specifically include Parkinson's  
CC disease, amyotrophic lateral sclerosis (ALS), Alzheimer's disease,  
CC stroke and various degenerative disorders affecting the brain.

XX Sequence 225 AA:

Query Match 100.0%; Score 1224; DB 19; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.6e-121;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMCLACTCTVLMHLPAPALNRTGDPGPGPSIQTKYDITRYLEHOLRLSLAGT 60  
DB 1 mdlragsdswgmclactctvlmhpavpalnrtgdpgpsiqtkydltrylehqlrslagtl 60  
QY 61 YLNYLGPPEPNEPDPNPRIGAEETLPRAVNLKRWRSINDRLRLTONYEAVSHILCTLRGL 120  
DB 61 ylnylgppnepdpnprigaeetlpratvnlkwrslndrlrltonyeavshilcylrql 120  
QY 121 NRQATRAELRSLAHFCTSLQGLIGSTAGYMATLGYPLDPLPTEPMAWPGAHSDFLQ 180  
DB 121 nrqaataelrslahfctslqgligstagymatlgyplpplptepmawpgahsdflq 180  
QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHEAHGF 225  
DB 181 kmddfwlkeltlwlrskdfnrlkkkmpaasvvlheahgf 225

RESULT 3

AAW87814  
ID AAW87814 standard; Protein: 225 AA.

XX AAW87814;

XX 24-AUG-2000 (first entry)

XX Murine NNT-1 protein.

XX NNT-1; neurotrophic factor; nootropic; neuroprotective; treatment;  
KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;  
KW nervous system degeneration; Alzheimer's disease; Parkinson's disease;  
KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome; murine;

KM Huntington's disease; peripheral neuropathy; neural retina degeneration;  
 KM retinopathy; immune disorder; hematopoietic disorder.  
 OS Mus sp.  
 XX US6054294-A.  
 XX 25-APR-2000.  
 PD 12-DEC-1997; 97US-098819.  
 XX 03-FEB-1997; 97US-0792019.  
 PR (AMGE-) AMGEN INC.  
 PA Chang M;  
 PI WPI: 2000-338492/29.  
 DR N-PSDB; AAA39483.  
 XX  
 PT New nucleic acids encoding neurotrophic factors useful for stimulating  
 PT growth of motor or sympathetic neurons for treating neuron cell damage  
 PT  
 PS  
 XX  
 XX Claim 2b: Fig 5; 42pp: English.  
 CC This invention describes a novel nucleic acid molecule (I) encoding a  
 CC novel neurotrophic factor (NNT-1) (II) which has nootropic,  
 CC neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and  
 CC ophthalmological activity. (I) is useful for producing NNT-1  
 CC polypeptides which are useful for treating patients in whom various  
 CC cells of the central, autonomic, or peripheral nervous system have  
 CC degenerated and/or have been damaged by congenital disease, trauma,  
 CC mechanical damage, surgery, stroke, ischemia, infection, metabolic  
 CC disease, nutritional deficiency, malignancy and/or toxic agents. NNT-1  
 CC proteoglycans are used to treat diseases like Alzheimer's, Parkinson's,  
 CC amyotrophic lateral sclerosis, Charcot-Marie-Tooth syndrome, Huntington's  
 CC disease, peripheral neuropathy induced by diabetes or other metabolic  
 CC disorders, and/or dystrophies or degeneration of the neural retina such  
 CC as retinitis pigmentosa, drug-induced retinopathies, stationary forms of  
 CC night blindness, progressive cone-rod degeneration, immune disorders and  
 CC hematopoietic disorders. (I) is effective in treating neurological  
 CC conditions and promotes neuron regeneration. Neural functions are  
 CC effectively restored in patients suffering from various neurological  
 CC disorders. This sequence represents the murine NNT-1 protein described in  
 CC the method of the invention.  
 CC  
 XX  
 XX Sequence 225 AA:  
 SQ  
 Query Match 100.0%; Score 1224; DB 21; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-121;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLRAGDSWGLACTCTVLMHLPVAPVLPALNRTGDPGPGSIQKTYDLTRYLEHQLSLAGT 60  
 Db 1 mdragswglactctvlmhlpavpavlnrtgdpqpgpslqkydltrylehqlslagt 60  
 QY 61 YLNLGPPFPPDFNPRLGAEVLPRAVTVLEWRSNLDRLRLTQNEAYSHLLCYRGL 120  
 Db 61 ylnlgppfppdfnprrlgaevlpravtvlewrsnldrlrltqneyashllcyrql 120  
 QY 121 NROAATLEIRSLAHCTSLQGLIGTAGVATGYRPLPOPLPCTEPMAWAGPASHDPLQ 180  
 Db 121 nrgataelirslahctslqgligtagvmatlgyrplpglptepawagpashdplq 180  
 QY 181 KMDPFWLKELOTWLMRSADKDFNRLKKMKPPAASVTLHLFAHGF 225  
 Db 181 kmdfwllkeltwlmrsakdfnrlkkmpaasvtlhleahgf 225  
 RESULT 4  
 AAM29715

ID AAM29715 standard; Protein; 225 AA.  
 XX  
 AC AAM29715;  
 XX  
 DT 09-NOV-1998 (first entry)  
 XX  
 DE Human neurotrophic factor NNT-1.  
 XX  
 KW NNT-1; neurotrophic factor; human; antiinflammatory; adjuvant;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;  
 KW peripheral neuropathy; dystrophy; neural retina degeneration;  
 KW common variable immunodeficiency; CVID; selective IgA deficiency;  
 KW hypogammaglobulinemia; X-linked agammaglobulinemia; antiseptic;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide 1..27  
 FT /label= Sig-peptide  
 FT 28..225  
 FT Protein /label= Mat-protein  
 FT  
 XX  
 XX W09833922-A1.  
 XX  
 PD 06-AUG-1998.  
 XX  
 PF 02-FEB-1998; 98WO-US02363.  
 XX  
 PR 30-JAN-1998; 98US-0016534.  
 PR 03-FEB-1997; 97US-0792019.  
 XX  
 PA (AMGE-) AMGEN INC.  
 PI Chang M, Elliott GS, Sarmiento U, Senaldi G;  
 DR WPI: 1998-437475/37.  
 DR N-PSDB: AAV47510-11.  
 XX  
 PT Newly isolated nucleic acid encoding human or murine neurotrophic  
 PT factor NNT-1 - useful for treatment of neurological and  
 PT immunological diseases or inflammation, also as vaccine adjuvant  
 PT  
 XX  
 PS Claim 12; Fig 3; 120pp: English.  
 XX  
 CC This is the amino acid sequence of a novel neurotrophic factor,  
 CC designated NNT-1, that is a growth factor for neurons in B or  
 CC T cells. It was deduced from isolated cDNA (see AAV47510) and  
 CC genomic DNA (see AAV47511) clones. Vectors containing the cDNA or  
 CC genomic DNA and host cells are provided for use in the production  
 CC of NNT-1 polypeptides. These are used to treat: (i) neurological  
 CC or immunological diseases, specifically Alzheimer's, Parkinson's  
 CC or Huntington's diseases, amyotrophic lateral sclerosis,  
 CC Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy and  
 CC degeneration of the neural retina, or conditions characterised by T  
 CC or B cell defects, e.g. common variable immunodeficiency (CVID),  
 CC selective IgA deficiency, hypogammaglobulinemia and X-linked  
 CC agammaglobulinemia (claimed), but many others disclosed; and (ii)  
 CC inflammation. NNT-1 is also able to boost immunoreactivity and  
 CC antibody production following vaccination, and, since it inhibits  
 CC tumour necrosis factor production, it may also be useful for  
 CC treating sepsis. In addition, cells that have been engineered to  
 CC express NNT-1 can be implanted, or nucleic acids are delivered in  
 CC gene therapy vectors.  
 CC  
 XX  
 SQ Sequence 225 AA:  
 Query Match 97.5%; Score 1193; DB 19; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 3e-118;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MDLRAGDSMGMLACCTVLMHLPAPVAPALNRTGDPGPGPSIQKTYDTLRYLEHQRLSAGT 60
DQ 1 mdlragswgmllacclctvlwhlpavpalnrtgdpqpgpslqktydlrlylehnqrlsagt 60
DQ 61 YLNYLGGPFNEPDNPRLGAETLPRAVTNLEWVRSNDRLRLTONFEAYSHLLCYLRGL 120
DQ 61 ylnylgpfnepdnprlgaetlpravtnlewvrsndrlrltonfeayshllcylrgl 120
QY 121 NRQAATAEPLRSLAHFCTSLQGLLGSLAGVMAATGYPPLPGLPCTEPAPAMPAPASDFLQ 180
DQ 121 nrgaataeplrslahfctslqglilgslagvmaatgypplpglpctepapampapahsdfllq 180
QY 181 KMDFWLKELQETWLMRSADFNRLKKMKQPPAASVTTLHLEAHGF 225
DQ 181 kmddfwlkeltqetwlmrsakdfnrlkkmkppaasvttlhlehagf 225

RESULT 5
AAW56141
ID AAW56141 standard; Protein: 225 AA.
AC AAW56141;
XX
XX
DT 13-JUL-1998 (first entry)
XX
DE Amino acid sequence of human neurotrophic factor NNT-1.
XX
XX Human: neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
XX treatment; neurological disease; degeneration; Parkinson's disease;
XX amyotrophic lateral sclerosis; ALS; Alzheimer's disease; stroke.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..27
XX FT Protein /note="signal peptide"
XX FT Protein 28..225 /note="mature protein"
XX
XX US5741772-A.
XX
XX 21-APR-1998.
XX
XX 03-FEB-1997; 97US-0792019.
XX
XX 03-FEB-1997; 97US-0792019.
XX
XX (AMGE-) AMGEN INC.
XX
XX Chang M;
XX
XX WPI: 1998-260526/23.
XX
XX N-PSDB: AAV22652.
XX
XX Neurotrophic factor NNT-1 polypeptide and related nucleic acids -
XX useful for stimulating growth of motor and sympathetic neurons
XX
XX Claim 1; Fig 3; 41pp: English.
XX
XX The present sequence represents a human neurotrophic factor, designated
XX NNT-1, which is capable of stimulating growth of motor or sympathetic
XX neurons. The NNT-1 protein is useful in the treatment of neurological
XX diseases characterized by the degeneration and death of particular
XX classes of neurons. These diseases specifically include Parkinson's
XX disease, amyotrophic lateral sclerosis (ALS), Alzheimer's disease,
XX stroke and various degenerative disorders affecting vision.
XX
XX Sequence 225 AA.

```

Query Match 97.5%; Score 1193; DB 19; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 36-118;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MDLRAGDSMGMLACCTVLMHLPAPVAPALNRTGDPGPGPSIQKTYDTLRYLEHQRLSAGT 60
DQ 1 mdlragswgmllacclctvlwhlpavpalnrtgdpqpgpslqktydlrlylehnqrlsagt 60
DQ 61 YLNYLGGPFNEPDNPRLGAETLPRAVTNLEWVRSNDRLRLTONFEAYSHLLCYLRGL 120
DQ 61 ylnylgpfnepdnprlgaetlpravtnlewvrsndrlrltonfeayshllcylrgl 120
QY 121 NRQAATAEPLRSLAHFCTSLQGLLGSLAGVMAATGYPPLPGLPCTEPAPAMPAPASDFLQ 180
DQ 121 nrgaataeplrslahfctslqglilgslagvmaatgypplpglpctepapampapahsdfllq 180
QY 181 KMDFWLKELQETWLMRSADFNRLKKMKQPPAASVTTLHLEAHGF 225
DQ 181 kmddfwlkeltqetwlmrsakdfnrlkkmkppaasvttlhlehagf 225

```

```

RESULT 6
AAW94466
ID AAW94466 standard; Protein: 225 AA.
AC AAW94466;
XX
XX
XX 22-APR-1999 (first entry)
XX
XX Human cardiotrophin-like cytokine protein.
XX
XX Human: cardiotrophin-like cytokine; interleukin 6 cytokine family;
XX CLC; IL-6; diagnosis; detection; immune system-related disorder;
XX cancer; cardiac disorder; heart failure; hypertension; cancer;
XX autoimmune disorder; infection.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..27
XX FT Protein /label= signal
XX FT Protein 28..225 /label= Cardiotrophin-like_cytokine
XX FT Domain 74..79 /label= CD-I
XX FT Domain /note="conserved domain"
XX FT Domain 150..156 /label= CD-II
XX FT Domain /note="conserved domain"
XX FT Domain 194..198 /label= CD-III
XX FT Domain /label= "conserved domain"
XX
XX WO990415-A1.
XX
XX 07-JAN-1999.
XX
XX 29-JUN-1998; 98WO-US13129.
XX
XX 30-JUN-1997; 97US-0051311.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Shi Y;
XX
XX WPI: 1999-095678/08.
XX
XX N-PSDB: AAX16161.
XX
XX New isolated cardiotrophin-like cytokine nucleic acid - used to
XX develop products for treating cardiac and immune system disorders,
XX e.g. heart failure, hypertension, cancers, autoimmune disorders and
XX infections
XX
XX Claim 1; Fig 1; 103pp: English.
XX
XX The present invention relates to a novel cardiotrophin-like cytokine

```



CC (CLC) protein which is a member of the interleukin 6 (IL-6) cytokine  
 CC family. The present sequence represents the human CLC protein. The  
 CC present invention also describes screening methods for identifying  
 CC agonists and antagonists of CLC activity, as well as methods for  
 CC detecting cardiac and immune system-related disorders and  
 CC therapeutic methods for treating cardiac and immune system-related  
 CC disorders, e.g. heart failure, hypertension, cancers, autoimmune  
 CC disorders and infections.  
 XX  
 XX Sequence 225 AA:

Query Match 97.5%; Score 1193; DB 20; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 3e-118;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMGLACTCTVLMHLPAPALNRTGPGGPGSTQKTYDITRYLHQLSLAGT 60  
 Db 1 mdlragswgmglactctvmlhlpapalnrtgpgpgstqktyditylqhlslag 60  
 QY 61 YLNYLGPPNEDEPNPRUGAETLPRAVNLKRWYSLNDRRLTQWYEAVALCYLRGL 120  
 Db 61 ylnylgppnefedpnprugaetlpravnlkrlwyslnndrlltqwyeavshlcyllrgl 120  
 QY 121 NROAATAEELRSLAHFCTSLQGLGSIAGVMTLGYPLPQPLPGTEPAPAPGPHSDFLQ 180  
 Db 121 nroaataelrrslahfctslqglgslagvmaalyplpqpplpgteptpaphsdflq 180  
 QY 181 KMDDFWLKLELQTWLRSKADFNRLKKMQPPAAVTLHLGAHF 225  
 Db 181 kmddfllkeltwtwlrskadfnrlkkmqppaavtlhlghaf 225

RESULT 7  
 AAY87813  
 ID AAY87813 standard; Protein; 225 AA.

XX AAY87813;

XX 24-AUG-2000 (first entry)

XX Human NNT-1 protein.

XX NNT-1, human; neurotrophic factor; neurotrophic; neuroprotective; treatment;  
 KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;  
 KW nervous system degeneration; Alzheimer's disease; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;  
 KW Huntington's disease; peripheral neuropathy; neural retina degeneration;  
 KW retinopathy; immune disorder; hematopoietic disorder.  
 XX  
 XX Homo sapiens.

XX US6054294-A.

XX 25-APR-2000.

XX 12-DEC-1997; 97US-0988819.

XX 03-FEB-1997; 97US-0792019.

XX (AMGE-) AMGEN INC.

XX Chang M;

XX WPI: 2000-338492/29.

XX N-PSDB; AAA39481.

PT New nucleic acids encoding neurotrophic factors useful for stimulating  
 PT growth of motor or sympathetic neurons for treating neuron cell damage  
 XX  
 XX Claim 1c: Fig 3; 42pp; English.

CC This invention describes a novel nucleic acid molecule (i) encoding a  
 CC novel neurotrophic factor (NNT-1) (II) which has neurotrophic,  
 CC neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and  
 CC ophthalmological activity. (I) is useful for producing NNT-1  
 CC polypeptides which are useful for treating patients in whom various  
 CC cells of the central, autonomic, or peripheral nervous system have  
 CC degenerated and/or have been damaged by congenital disease, trauma,  
 CC mechanical damage, surgery, stroke, ischemia, infection, metabolic  
 CC disease, nutritional deficiency, malignancy and/or toxic agents. NNT-1  
 CC proteins are used to treat diseases like Alzheimer's, Parkinson's,  
 CC amyotrophic lateral sclerosis, Charcot-Marie-Tooth syndrome, Huntington's  
 CC disease, peripheral neuropathy induced by diabetes or other metabolic  
 CC disorders, and/or dystrophies or degeneration of the neural retina such  
 CC as retinitis pigmentosa, drug-induced retinopathies, stationary forms of  
 CC night blindness, progressive cone-rod degeneration, immune disorders and  
 CC hematopoietic disorders. (I) is effective in treating neurological  
 CC conditions and promotes neuron regeneration. Neural functions are  
 CC effectively restored in patients suffering from various neurological  
 CC disorders. This sequence represents the human NNT-1 protein described in  
 CC the method of the invention.  
 XX  
 XX Sequence 225 AA:

Query Match 97.5%; Score 1193; DB 21; Length 225;  
 Best Local Similarity 96.9%; Pred. No. 3e-118;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMGLACTCTVLMHLPAPALNRTGPGGPGSTQKTYDITRYLHQLSLAGT 60  
 Db 1 mdlragswgmglactctvmlhlpapalnrtgpgpgstqktyditylqhlslag 60  
 QY 61 YLNYLGPPNEDEPNPRUGAETLPRAVNLKRWYSLNDRRLTQWYEAVALCYLRGL 120  
 Db 61 ylnylgppnefedpnprugaetlpravnlkrlwyslnndrlltqwyeavshlcyllrgl 120  
 QY 121 NROAATAEELRSLAHFCTSLQGLGSIAGVMTLGYPLPQPLPGTEPAPAPGPHSDFLQ 180  
 Db 121 nroaataelrrslahfctslqglgslagvmaalyplpqpplpgteptpaphsdflq 180  
 QY 181 KMDDFWLKLELQTWLRSKADFNRLKKMQPPAAVTLHLGAHF 225  
 Db 181 kmddfllkeltwtwlrskadfnrlkkmqppaavtlhlghaf 225

RESULT 8

AA663543  
 ID AA663543 standard; Protein; 225 AA.

XX AA663543;

XX 15-OCT-2001 (first entry)

XX Amino acid sequence of a human NNT-1 protein.

XX NNT-1, CLF-1, SCNTFALPHA; nervous system; neuron; nervous system;  
 KW neuro-muscular function; tumour; immune system; hematopoietic system;  
 KW reproductive system; liver; skeletal muscle; neurodegenerative disease;  
 KW amyotrophic lateral sclerosis; Parkinson's disease; Huntington's disease;  
 KW muscular mass; paralysis; cancer; obesity; fertility; endometriosis;  
 KW blastocyst implantation; thrombosis; retinal disease;  
 KW retinal pigmentosis.  
 XX  
 XX Homo sapiens.

XX WO200155172-A2.

XX 02-AUG-2001.

XX 26-JAN-2001; 2001WO-FR00253.

XX 27-JAN-2000; 2000FR-0001035.

XX 12-OCT-2000; 2000FR-0013089.



Db 89 ylnylgppfnepdnpprlgaetlprativdewrslncklrltqnyeaashllylrgl 148  
 QY 121 NRQAATRELKSLAHFCTSLQGLIGSIAGVWATLGYPLPQPLPCTEPAMAGPAHSDFLQ 180  
 Db 149 nrgataeltrslahfctslgllgsiagymaalgyrlpglpptetpwrpghsdrlq 208  
 QY 181 KMDFWLKELQTLWMSAKDFNLKKMKQPPAASVTLHEANGF 225  
 Db 209 kmddfllkeltqlwrsakdfnrlkkmqppaaavclhlgahf 253

## RESULT 10

AAE00828

ID AAE00828 standard; Protein; 223 AA.

XX AAE00828;

DT 02-JUL-2001 (first entry)

XX Human cardiotrophin-like cytokine (CLC) protein.

XX Human: biologically active complex; haemopoietin receptor; NR6;  
 KW cardiotrophin-like cytokine; CLC; therapy; prophylaxis; proliferation;  
 XX differentiation; cell survival; neurotrophic activity.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..27

FT Protein /label= Signal\_peptide

FT /label= Human\_mature\_CLC\_protein

FT /note= "Cardiotrophin-like cytokine"

XX MO200127157-A1.

XX 19-APR-2001.

XX 06-OCT-2000; 2000WO-AU01216.

XX 08-OCT-1999; 99AU-0003327.

XX 12-MAY-2000; 2000AU-0007489.

XX (AMRA-) AMRAD OPERATIONS PTY LTD.

XX PI Nash A, Jachno KM, Fabri LJ, Reid K, Bartlett PF, Hilton DJ;

XX PI Nakata Y, Hasegawa M;

XX DR N-PSDB; AAD04201.

XX WPI; 2001-281978/29.

XX New biologically active complex comprising NR6 and  
 PT cardiotrophin-like cytokine, for facilitating proliferation,  
 PT differentiation and/or survival of a cell

XX Claim 32; Page 114-115; 123pp; English.

XX The present invention relates to a biologically active complex comprising  
 CC a haemopoietin receptor, NR6 and cardiotrophin-like cytokine (CLC).  
 CC The complex is useful in the manufacture of a medicament for the  
 CC treatment and/or prophylaxis of a subject, as it is involved in  
 CC facilitating proliferation, differentiation and/or survival of a cell.  
 CC The complex or its components have neurotrophic activity. The present  
 CC sequence is human cardiotrophin-like cytokine (CLC) protein.

XX Sequence 223 AA;

## Query Match

Best Local Similarity 96.5%; Score 1181; DB 22;

Matches 216; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMACCTVLMHLPAVPALNRGDPGPGSIQKTYDILFYLEHJLHSLAGT 60  
 Db 1 mdlragdswgmactclvhlpavpalnrtgdpdpgpslqktydlfylehlslslagt 60  
 QY 61 YLNLGPPFNEPDNPPLGAETLPRAVNLQVWMSLNDRLRLQNTYAYSHLQCYLRGL 120  
 Db 61 ylnylgppfnepdnpprlgaetlprativdewrslncklrltqnyeaashllylrgl 120  
 QY 121 NRQAATRELKSLAHFCTSLQGLIGSIAGVWATLGYPLPQPLPCTEPAMAGPAHSDFLQ 180  
 Db 121 nrgataeltrslahfctslgllgsiagymaalgyrlpglpptetpwrpghsdrlq 180  
 QY 181 KMDFWLKELQTLWMSAKDFNLKKMKQPPAASVTLHEANGF 223  
 Db 181 kmddfllkeltqlwrsakdfnrlkkmqppaaavclhlgahf 223

## RESULT 11

AAB19587

ID AAB19587 standard; Protein; 215 AA.

XX AAB19587;

DT 22-JAN-2001 (first entry)

XX Mouse Interleukin-B60 (IL-B60).

XX Interleukin-B60; IL-B60; mouse; cytokine; cytokine-1; factor-1;  
 KW haematopoietic; inflammation; antiinflammatory; autoimmune disease;  
 XX therapy.

XX Mus musculus.

XX

FH Key Location/Qualifiers

FT Peptide 1..17

FT Protein /label= Signal\_peptide

FT /label= Mature-protein

XX WO200053631-A1.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06182.

XX 11-MAR-1999; 99US-0267901.

XX (SCHE) SCHERING CORP.

XX Oppmann B, Timans JC, Kastelein RA, Bazan JF;

XX WPI; 2000-587426/55.

XX N-PSDB; AAA88547.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Claim 1; Page 17; 97pp; English.

XX The present sequence is that of mouse Interleukin-B60 (IL-B60), a  
 CC novel, small soluble cytokine-like protein that exhibits structural  
 CC motifs characteristic of a member of the long-chain cytokines, and  
 CC which shows homology to granulocyte colony stimulating factor and  
 CC interleukin-6. IL-B60 may have either stimulatory or inhibitory  
 CC effects on haematopoietic cells, including e.g. lymphoid cells,  
 CC such as T-cells, B-cells, natural killer cells, macrophages,  
 CC dendritic cells, haematopoietic progenitors, etc. Methods are  
 CC provided for modulating the physiology or development of a cell or  
 CC tissue culture cells by contacting the cell with an agonist or  
 CC antagonist of IL-B60 or an agonist of antagonist of a complex of  
 CC mature IL-B60 and its partner, cytokine-like factor-1 (CLF-1, see  
 CC AAB19588). The IL-B60/CLF-1 cytokine serves as a key physiological

CC factor in motor neuron development and regeneration. IL-60B, its  
CC agonists and antagonists may be used to treat inflammatory or  
CC autoimmune disorders and also for drug screening.

Sequence 215 AA:

Query Match 95.3%; Score 1167; DB 21; Length 215;  
Best Local Similarity 100.0%; Pred. No. 1.6e-115;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 MLaCTCTVWHLPAVPALNRTGDPGPGPSIQKTYDLYRLEHOLRSAGTYVNLGPPFN 70  
DB 1 mLaclctvlwhlpavpalnrtgdpdpqpsiqktydlyrlehqrlsagtylnylgppfn 60  
QY 71 EPPFNPRPGAETLPRAVNLQVWRSINDRLRLTQNEEAYSHLLCYLGLNROAATAELR 130  
DB 61 epdlnpprlgaetlpratlvdlewrslndrlrltqneyaayshllcylrglnrgataelr 120  
QY 131 RSLAHFCTSLQGLGSLGAGVMAATLGYPLPQPLPGTEPAMAPGAHSDPFLQKMDFWLLKE 190  
DB 121 rslahfctslqglgslgagvmaatlgyplpqplpgtepamapghasdpflqkmdfwllke 180  
QY 191 LQTWLMSAKDFNRLKKMKOPPAASVTLHLEAHGF 225  
DB 181 lqtwlmsakdfnrlkkmpaasvtlhlea hgf 215

# RESULT 12

AAB19586 standard; Protein; 215 AA.

AC AAB19586;

DF 22-JAN-2001 (first entry)

DE Human Interleukin-B60 (IL-B60).

XX Interleukin-B60; IL-B60; human; cytokine; cytokine-like factor-1;  
KW haematopoietic; inflammatory; antinflammatory; autoimmune disease;  
therapy.

OS Homo sapiens.

XX Key Location/Qualifiers

FF Peptide 1..17

FF Protein /Label= Signal\_peptide

FF Protein /Label= Mature-protein

FN WO200053631-A1.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06182.

XX 11-MAR-1999; 99US-0267901.

XX (SCHE) SCHERING CORP.

PI Opmann B, Timans JC, Kastelein RA, Bazan JF;

DR N-PSDB; AAA88546.

XX Cytokine-like factor 1 (CLF-1) and interleukin (IL)-B60 complexes,  
PT polypeptides, and nucleic acids, useful in research, diagnosis and for  
PT treating inflammatory and autoimmune disorders

PS Claim 1; Page 15-16; 97pp; English.

XX The present sequence is that of human interleukin-B60 (IL-B60), a  
CC novel, small soluble cytokine-like protein that exhibits structural

CC motifs characteristic of a member of the long-chain cytokines, and  
CC which shows homology to granulocyte colony stimulating factor and  
CC interleukin-6. IL-60B may have either stimulatory or inhibitory  
CC effects on haematopoietic cells, including e.g. lymphoid cells,  
CC such as T-cells, B-cells, natural killer cells, macrophages,  
CC dendritic cells, haematopoietic progenitors, etc. Methods are  
CC provided for modulating the physiology or development of a cell or  
CC tissue culture cells by contacting the cell with an agonist or  
CC antagonist of IL-B60 or an agonist of antagonist of a complex of  
CC mature IL-B60 and its partner, cytokine-like factor-1 (CLF-1, see  
CC AAB19586). The IL-B60/CLF-1 cytokine serves as a key physiological  
CC factor in motor neuron development and regeneration. IL-60B, its  
CC agonists and antagonists may be used to treat inflammatory or  
CC autoimmune disorders and also for drug screening.

Sequence 215 AA:

Query Match 92.8%; Score 1136; DB 21; Length 215;  
Best Local Similarity 96.7%; Pred. No. 3.1e-112;  
Matches 208; Conservative 3; Mismatches 4; Indels 1; Gaps 0;

QY 11 MLaCTCTVWHLPAVPALNRTGDPGPGPSIQKTYDLYRLEHOLRSAGTYVNLGPPFN 70  
DB 1 mLaclctvlwhlpavpalnrtgdpdpqpsiqktydlyrlehqrlsagtylnylgppfn 60  
QY 71 EPPFNPRPGAETLPRAVNLQVWRSINDRLRLTQNEEAYSHLLCYLGLNROAATAELR 130  
DB 61 epdlnpprlgaetlpratlvdlewrslndrlrltqneyaayshllcylrglnrgataelr 120  
QY 131 RSLAHFCTSLQGLGSLGAGVMAATLGYPLPQPLPGTEPAMAPGAHSDPFLQKMDFWLLKE 190  
DB 121 rslahfctslqglgslgagvmaatlgyplpqplpgtepamapghasdpflqkmdfwllke 180  
QY 191 LQTWLMSAKDFNRLKKMKOPPAASVTLHLEAHGF 225  
DB 181 lqtwlmsakdfnrlkkmpaasvtlhlea hgf 215

# RESULT 13

AAM20115 standard; Protein; 164 AA.

AC AAM20115;

DF 12-OCT-2001 (first entry)

DE Peptide #6549 encoded by probe for measuring cervical gene expression.

DE Probe: human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer.

OS Homo sapiens.

XX WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SC, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX

PS Claim 27; SEQ ID No 24941; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 164 AA:

Query Match 69.6%; Score 852; DB 22; Length 164;  
Best Local Similarity 95.1%; Pred. No. 2.7e-82;  
Matches 156; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 62 LNYLGPPNEPDPNPRLGAETLPRAVNLVWRSNDRLRLTONYAYSHLLCYLRGLN 121  
Db 1 lnylgppnepdpnpriqaktlpratlvdlewrslndkrlrlqnyeyshllcylrqln 60

OY 122 RQAATAELRSLAHFCTSLGGLGSLAGVMAATLGYPLPPLPGTEPAPAPGASDFLQK 181  
Db 61 rqaataelrslahfctslgglgslagvmaalgyplpplpgtepwprgahsdflqk 120

OY 182 MDDEFWLKELQTLWMSAKDFNRLKKMKOPAPASVTLHLEAHGF 225  
Db 121 mddfwwlkelqtlwmsakdfnrlkkmkppaavtlhlgahgf 164

RESULT 14

AA034012 standard; Protein; 164 AA.

AC AAM34012;

DT 17-OCT-2001 (first entry)

DE Peptide #8049 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

OS Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000US-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX

PS Claim 27; SEQ ID No 34281; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;  
CC see AAI3315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX Sequence 164 AA:

Query Match 69.6%; Score 852; DB 22; Length 164;  
Best Local Similarity 95.1%; Pred. No. 2.7e-82;  
Matches 156; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 62 LNYLGPPNEPDPNPRLGAETLPRAVNLVWRSNDRLRLTONYAYSHLLCYLRGLN 121  
Db 1 lnylgppnepdpnpriqaktlpratlvdlewrslndkrlrlqnyeyshllcylrqln 60

OY 122 RQAATAELRSLAHFCTSLGGLGSLAGVMAATLGYPLPPLPGTEPAPAPGASDFLQK 181  
Db 61 rqaataelrslahfctslgglgslagvmaalgyplpplpgtepwprgahsdflqk 120

OY 182 MDDEFWLKELQTLWMSAKDFNRLKKMKOPAPASVTLHLEAHGF 225  
Db 121 mddfwwlkelqtlwmsakdfnrlkkmkppaavtlhlgahgf 164

RESULT 15

AA09197 standard; Protein; 208 AA.

AC AAY09197;

DT 20-JUL-1999 (first entry)

DE Human DNAX interleukin-40 (DIL-40) alternative sequence.

KW DNAX interleukin-40; DIL-40; activation regulation; development;  
KW differentiation; hematopoietic; neural cell; drug targeting;  
KW immunological condition.

OS Homo sapiens.

PN W09919491-A2.

PD 22-APR-1999.

PF 13-OCT-1998; 98WO-US20932.

PR 14-OCT-1997; 97US-0950238.

PR (SCHE ) SCHERING CORP.

PR Bazan JF, Kastelein RA;

PR WPI; 1999-312475/26.

PR N-PSDB; AAX60797.

PR New human DNAX interleukin-40 (DIL-40) polypeptide and mutants

PS Claim 1; Page 10-11; 76pp; English.

XX The invention relates to a human DNAX interleukin-40 (DIL-40) protein.  
CC Host cells containing a vector comprising the DIL-40 nucleic acid can be  
CC used for the recombinant expression of the protein. The DIL-40 protein,  
CC gene and antibody are useful for regulating activation, development  
CC differentiation and function of various cell types, including  
CC hematopoietic or neural cells. The antibodies may be anti-DIL-40, anti-DIL-40 mutants, and



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:31:19 ; Search time 30.45 Seconds  
(without alignments)  
166,280 Million cell updates/sec

Title: US-09-931-704-5

Perfect score: 1224  
Sequence: 1 MDLRAGDSWGMCLCTVLM.....KKMKPPASVYLHLEHGF 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/CTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1224	100.0	225	1	US-08-792-019B-5
2	1224	100.0	225	3	US-08-988-819-5
3	1224	100.0	225	4	US-09-016-534-5
4	1193	97.5	225	1	US-08-792-019B-2
5	1193	97.5	225	3	US-09-106-182-2
6	1193	97.5	225	3	US-08-988-819-2
7	1193	97.5	225	4	US-09-016-534-2
8	118.5	9.7	203	3	US-09-106-182-3
9	110.5	9.0	203	1	US-08-233-609-3
10	110.5	9.0	203	1	US-08-444-083-3
11	110.5	9.0	203	1	US-08-286-304-3
12	110.5	9.0	203	1	US-08-442-745-3
13	110.5	9.0	203	1	US-08-443-129-3
14	110.5	9.0	203	1	US-08-443-952-3
15	110.5	9.0	203	1	US-08-443-130-3
16	110.5	9.0	203	3	US-08-898-911-3
17	110.5	9.0	203	5	PCT-US95-04467-3
18	91.5	7.5	201	1	US-08-444-083-8
19	91.5	7.5	201	1	US-08-286-304-8
20	91.5	7.5	201	1	US-08-442-745-8
21	91.5	7.5	201	1	US-08-443-129-8
22	91.5	7.5	201	1	US-08-443-952-8
23	91.5	7.5	201	1	US-08-443-130-8
24	91.5	7.5	201	1	US-08-792-019B-11
25	91.5	7.5	201	3	US-09-106-182-4
26	91.5	7.5	201	3	US-08-988-819-11
27	91.5	7.5	201	3	US-08-898-911-8

28	91.5	7.5	201	4	US-09-016-534-11	Sequence 11, Appl
29	91.5	7.5	201	5	PCT-US95-04467-8	Sequence 8, Appl
30	90	7.4	195	1	US-07-950-284-5	Sequence 5, Appl
31	90	7.4	195	2	US-08-308-736A-5	Sequence 5, Appl
32	90	7.4	195	5	PCT-US92-08258-2	Sequence 2, Appl
33	90	7.4	195	5	PCT-US93-09649A-5	Sequence 5, Appl
34	90	7.4	195	5	PCT-US93-09649-5	Sequence 5, Appl
35	85.5	7.0	560	2	US-08-756-317-11	Sequence 11, Appl
36	82.5	6.7	2628	4	US-09-413-814-11	Sequence 9, Appl
37	81.5	6.7	582	2	US-08-422-699A-9	Sequence 16, Appl
38	81.5	6.7	582	2	US-08-422-706B-9	Sequence 4, Appl
39	80.5	6.6	379	4	US-09-105-537-16	Sequence 2, Appl
40	80.5	6.6	378	4	US-09-105-537-4	Sequence 2, Appl
41	79	6.5	1620	1	US-08-542-363-2	Sequence 6, Appl
42	79	6.5	1620	4	US-09-100-089-2	Sequence 2, Appl
43	76.5	6.2	379	4	US-09-320-878-6	Sequence 6, Appl
44	76.5	6.2	388	2	US-08-759-581B-9	Sequence 2, Appl
45	75.5	6.2	593	3	US-08-964-268-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-792-019B-5  
: Sequence 5, Application US/08792019B  
: Patent No. 5741772  
: GENERAL INFORMATION:  
: APPLICANT: CHANG, MING-SHI  
: TITLE OF INVENTION: THE NEUROTROPIC FACTOR NNT-1  
: NUMBER OF SEQUENCES: 16  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: AMGEN INC.  
: STREET: 1840 DEHAVILLAND DRIVE  
: CITY: THOUSAND OAKS  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 91320  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/792,019B  
: FILING DATE: 03-FEB-1997  
: CLASSIFICATION: 514  
: ATTORNEY/AGENT INFORMATION:  
: NAME: COOK, ROBERT R. 31,602  
: REGISTRATION/DOCKET NUMBER: A-442  
: INFORMATION FOR SEQ ID NO: 5:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 225 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-08-792-019B-5

Query Match 100.0%; Score 1224; DB 1; Length 225;  
Best Local Similarity 100.0%; Pred. No. 5.9e-122;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLRAGDSWGMCLCTVLMHLPALNRGTGPGPSIQKTYDLTKLH LRSAGT 60  
DB 1 MDLRAGDSWGMCLCTVLMHLPALNRGTGPGPSIQKTYDLTKLH LRSAGT 60  
QY 61 YLVYLPPEPPEPFPNPRIGAEFLPRATVNLKWRSLNRLRLTONYEVNS LLYLRL 120  
DB 61 YLVYLPPEPPEPFPNPRIGAEFLPRATVNLKWRSLNRLRLTONYEVNS LLYLRL 120  
QY 121 NROAATLRRSLAHKCTSIQGLLGSISAGYMATLGYPLDQPLGTEPMAV AHSDFIQ 180

Db 121 NRQAATLRLSLAHFCTSLQGLGSIAGVMTLGYPLPQLPCTEPAMAPGPAHSDFLQ 180  
QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLEAHGF 225  
Db 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLEAHGF 225

## RESULT 2

US-08-988-819-5  
; Sequence 5, Application US/08988819  
; Patent No. 6054294

## GENERAL INFORMATION:

APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,819  
FILING DATE: 12-DEC-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442A  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-988-819-5

Query Match 100.0%; Score 1224; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 5.9e-122;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMCLACTCTVLMHLPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
Db 1 MDLRAGDSWGMCLACTCTVLMHLPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
QY 61 YLNYLGPPNEPDPNPRRLGAEITLPRAIVNLEWWSLNDRLRLTONYEAVSHLLCYLRGL 120  
Db 61 YLNYLGPPNEPDPNPRRLGAEITLPRAIVNLEWWSLNDRLRLTONYEAVSHLLCYLRGL 120  
QY 121 NRQAATLRLSLAHFCTSLQGLGSIAGVMTLGYPLPQLPCTEPAMAPGPAHSDFLQ 180  
Db 121 NRQAATLRLSLAHFCTSLQGLGSIAGVMTLGYPLPQLPCTEPAMAPGPAHSDFLQ 180  
QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLEAHGF 225  
Db 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLEAHGF 225

RESULT 3  
US-09-016-534-5  
; Sequence 5, Application US/09016534  
; Patent No. 6143874

GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
APPLICANT: ELLIOTT, GARY S.  
APPLICANT: SARMIENTO, ULLA  
APPLICANT: SEMALDI, GIORGIO  
TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: USA  
ZIP: 91320

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,534  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/792,019  
FILING DATE: 03-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-442B  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-016-534-5

Query Match 100.0%; Score 1224; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 5.9e-122;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGMCLACTCTVLMHLPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
Db 1 MDLRAGDSWGMCLACTCTVLMHLPALNRTGDPGPGSIQKTYDLTRYLEHQLRSLAGT 60  
QY 61 YLNYLGPPNEPDPNPRRLGAEITLPRAIVNLEWWSLNDRLRLTONYEAVSHLLCYLRGL 120  
Db 61 YLNYLGPPNEPDPNPRRLGAEITLPRAIVNLEWWSLNDRLRLTONYEAVSHLLCYLRGL 120  
QY 121 NRQAATLRLSLAHFCTSLQGLGSIAGVMTLGYPLPQLPCTEPAMAPGPAHSDFLQ 180  
Db 121 NRQAATLRLSLAHFCTSLQGLGSIAGVMTLGYPLPQLPCTEPAMAPGPAHSDFLQ 180  
QY 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLEAHGF 225  
Db 181 KMDDFWLKELQTLWLRSAKDFNRLKKKMPAASVTLHLEAHGF 225

RESULT 4  
US-08-792-019B-2  
; Sequence 2, Application US/08792019B  
; Patent No. 5741772  
GENERAL INFORMATION:  
APPLICANT: CHANG, MING-SHI  
TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: 1840 DEHAVILLAND DRIVE  
CITY: THOUSAND OAKS  
STATE: CA



COUNTRY: USA  
 ZIP: 91320  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/792,019B  
 FILING DATE: 03-FEB-1997  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: COOK, ROBERT R.  
 REGISTRATION NUMBER: 31,602  
 REFERENCE/DOCKET NUMBER: A-442  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 225 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-792-019B-2

Query Match  
 Best Local Similarity 97.5%; Score 1193; DB 1; Length 225;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGLACLTVMHLPAPVALNRTGDPGPGSIQKTYDTRYLEHQLRSLAGT 60  
 DB 1 MDLRAGDSWGLACLTVMHLPAPVALNRTGDPGPGSIQKTYDTRYLEHQLRSLAGT 60  
 QY 61 YLNYLGPFPNEPDPNPRLGAETLPRAIVNLEWRSINDRLRLTONYEAVSHLLCYLRGL 120  
 DB 61 YLNYLGPFPNEPDPNPRLGAETLPRAIVNLEWRSINDRLRLTONYEAVSHLLCYLRGL 120  
 QY 121 NROATAELRSLAHFCTSLQGLGSIAGVATLGYPLPQPLPGTEPAMAVLHNSDFLQ 180  
 DB 121 NROATAELRSLAHFCTSLQGLGSIAGVATLGYPLPQPLPGTEPAMAVLHNSDFLQ 180  
 QY 181 KMDPFWLKELOTWLMRSKAKDFNRLKKKMPAPASVTLHLEAHGF 225  
 DB 181 KMDPFWLKELOTWLMRSKAKDFNRLKKKMPAPASVTLHLEAHGF 225

RESULT 5  
 US-09-106-182-2  
 Sequence 2, Application US/09106182  
 Patent No. 6046035  
 GENERAL INFORMATION:  
 APPLICANT: Shi, Yangu  
 APPLICANT: Ruden, Steve  
 TITLE OF INVENTION: Cardiotrophin-Like Cytokine  
 NUMBER OF SEQUENCES: 24  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc  
 STREET: 9410 Key West Ave  
 CITY: Rockville  
 STATE: MD  
 COUNTRY: US  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/106,182  
 FILING DATE: Herewith  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/051,053  
 FILING DATE: 30-JUN-1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Brookes, A. Anders  
 REGISTRATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: PF885  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-309-8504  
 TELEFAX: 301-309-8439  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 225 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-106-182-2

Query Match  
 Best Local Similarity 97.5%; Score 1193; DB 3; Length 225;  
 Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLRAGDSWGLACLTVMHLPAPVALNRTGDPGPGSIQKTYDTRYLEHQLRSLAGT 60  
 DB 1 MDLRAGDSWGLACLTVMHLPAPVALNRTGDPGPGSIQKTYDTRYLEHQLRSLAGT 60  
 QY 61 YLNYLGPFPNEPDPNPRLGAETLPRAIVNLEWRSINDRLRLTONYEAVSHLLCYLRGL 120  
 DB 61 YLNYLGPFPNEPDPNPRLGAETLPRAIVNLEWRSINDRLRLTONYEAVSHLLCYLRGL 120  
 QY 121 NROATAELRSLAHFCTSLQGLGSIAGVATLGYPLPQPLPGTEPAMAVLHNSDFLQ 180  
 DB 121 NROATAELRSLAHFCTSLQGLGSIAGVATLGYPLPQPLPGTEPAMAVLHNSDFLQ 180  
 QY 181 KMDPFWLKELOTWLMRSKAKDFNRLKKKMPAPASVTLHLEAHGF 225  
 DB 181 KMDPFWLKELOTWLMRSKAKDFNRLKKKMPAPASVTLHLEAHGF 225

RESULT 6  
 US-08-988-819-2  
 Sequence 2, Application US/08988819  
 Patent No. 6054294  
 GENERAL INFORMATION:  
 APPLICANT: CHANG, MING-SHI  
 TITLE OF INVENTION: NEUTROTROPHIC FACTOR NNT-1  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: AMGEN INC.  
 STREET: ONE AMGEN CENTER DRIVE  
 CITY: THOUSAND OAKS  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 91320  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/988,819  
 FILING DATE: 12-DEC-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/792,019  
 FILING DATE: 03-FEB-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: COOK, ROBERT R.  
 REGISTRATION NUMBER: 31,602  
 REFERENCE/DOCKET NUMBER: A-442A  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 225 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear



OY 152 ATLCY-----PLPQPLPCTEPAPGAHSDFLQKMDPWLKELQOTWLMRSKADFNRL 205  
 141 AALGAARGVPEPV-ATSLFTSNSAGVFSKAVLGLHVGCLYGEWVSRTGDLGOL 197

RESULT 9

US-08-233-609-3  
 ; Sequence 3, Application US/08233609  
 ; Patent No. 5534615  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Joffre  
 ; APPLICANT: Chien, Kenneth  
 ; APPLICANT: King, Kathleen  
 ; APPLICANT: Pennica, Diane  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patin (Genentech)  
 ; CURRENT APPLICATION NUMBER: US/08/233,609  
 ; FILING DATE: 25-APR-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; FILING DATE:  
 ; NAME: Hasak, Janet E.  
 ; REGISTRATION NUMBER: 28,616  
 ; REFERENCE/DOCKET NUMBER: 894  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/952-1896  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 203 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-233-609-3

Query Match 9.0%; Score 110.5; DB 1; Length 203;  
 Best Local Similarity 27.4%; Pred. No. 0.0005;  
 Matches 49; Conservative 23; Mismatches 86; Indels 21; Gaps 6;

OY 40 IOKTYDITRYLEHQLRSAGTYLNYLGPPEDEFPNPRU---GAETLRATYVLEWRS 96  
 27 IROTHNLARLLTYAEQLLEBYVOOGEPGLPGFSPRLPLAGLSGPAVSHAGLPV--- 83  
 DB 97 LNDRLRLTONYEAASHLLCYLRGLNROA-----TALRSLAHFCSLQGLGSIAGVM 151  
 DB 84 -SERLR--QDAALSVLPALLDVARRQALNRPAPRLRSLSDAARQVRAVETVL 140  
 OY 152 ATLCYPLPQPLPCTEPA-----WAPGAHSDFLQKMDPWLKELQOTWLMRSKADFNRL 205  
 DB 141 AALGAARAG--PGPEPVYATVLEFTANSAGIFSAKVLGFHVGCLYGEWVSRTGDLGOL 197

RESULT 10  
 US-08-444-083-3  
 ; Sequence 3, Application US/08444083  
 ; Patent No. 5571675

GENERAL INFORMATION:  
 ; APPLICANT: Baker, Joffre  
 ; APPLICANT: Chien, Kenneth  
 ; APPLICANT: King, Kathleen  
 ; APPLICANT: Pennica, Diane  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/444,083  
 ; FILING DATE: 17-MAY-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/233609  
 ; FILING DATE: 25-APR-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/286304  
 ; FILING DATE: 05-AUG-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Torchia, Timothy E.  
 ; REGISTRATION NUMBER: 36,700  
 ; REFERENCE/DOCKET NUMBER: 894PIDS  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/952-8674  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 203 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-444-083-3

Query Match 9.0%; Score 110.5; DB 1; Length 203;  
 Best Local Similarity 27.4%; Pred. No. 0.0005;  
 Matches 49; Conservative 23; Mismatches 86; Indels 21; Gaps 6;

OY 40 IOKTYDITRYLEHQLRSAGTYLNYLGPPEDEFPNPRU---GAETLRATYVLEWRS 96  
 27 IROTHNLARLLTYAEQLLEBYVOOGEPGLPGFSPRLPLAGLSGPAVSHAGLPV--- 83  
 DB 97 LNDRLRLTONYEAASHLLCYLRGLNROA-----TALRSLAHFCSLQGLGSIAGVM 151  
 DB 84 -SERLR--QDAALSVLPALLDVARRQALNRPAPRLRSLSDAARQVRAVETVL 140  
 OY 152 ATLCYPLPQPLPCTEPA-----WAPGAHSDFLQKMDPWLKELQOTWLMRSKADFNRL 205  
 DB 141 AALGAARAG--PGPEPVYATVLEFTANSAGIFSAKVLGFHVGCLYGEWVSRTGDLGOL 197

RESULT 11  
 US-08-286-304-3  
 ; Sequence 3, Application US/08286304  
 ; Patent No. 5571893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Joffre  
 ; APPLICANT: Chien, Kenneth  
 ; APPLICANT: King, Kathleen  
 ; APPLICANT: Pennica, Diane  
 ; APPLICANT: Wood, William

TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,304  
FILING DATE: 05-AUG-1994  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/233609  
FILING DATE: 25-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 894P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 203 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-286-304-3

Query Match  
Best Local Similarity 27.4%; Score 110.5; DB 1; Length 203;  
Matches 49; Conservative 23; Mismatches 86; Indels 21; Gaps 6;

QY 40 IOKTYDLTRYLEHQLSLAGTYLNTLGPFPNEDPDPNPRLL---GAETLPRAVNLFWRS 96  
DB 27 IROHNLARLLTKYAEOLEEVYQGGPFGIPGSPRLPLAGISGAPSHAGLPV--- 83  
QY 97 LNDRLRLTQNTYEAASHLLCYLRGLNRQA---TAELRSLAHFCTSLUGLLSGIAGYM 151  
DB 84 -SERLR--QDAALSVLPALLDAVRRROELNPRAPRLRLSLSDAAROVRAAGAVEYVL 140  
QY 152 ATUGPPLPQPLPGTEPA-----WAPGPAHSDFLOKMDDFWLKELCTWLMRSKADENRL 205  
DB 141 AALGAARG--PGPEPVYATLTFTANSTAGIFSAYLGFHVGLGVEWVSRTFGDGLGOL 197

RESULT 12  
US-08-442-745-3  
Sequence 3, Application US/08442745  
Patent No. 5624806

GENERAL INFORMATION:  
APPLICANT: Baker, Joffre  
APPLICANT: Chien, Kenneth  
APPLICANT: King, Kathleen  
APPLICANT: Pennica, Diane  
APPLICANT: Wood, William  
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,745  
FILING DATE: 17-may-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/233609  
FILING DATE: 25-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: 894P1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8674  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 203 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-442-745-3

Query Match  
Best Local Similarity 27.4%; Score 110.5; DB 1; Length 203;  
Matches 49; Conservative 23; Mismatches 86; Indels 21; Gaps 6;

QY 40 IOKTYDLTRYLEHQLSLAGTYLNTLGPFPNEDPDPNPRLL---GAETLPRAVNLFWRS 96  
DB 27 IROHNLARLLTKYAEOLEEVYQGGPFGIPGSPRLPLAGISGAPSHAGLPV--- 83  
QY 97 LNDRLRLTQNTYEAASHLLCYLRGLNRQA---TAELRSLAHFCTSLUGLLSGIAGYM 151  
DB 84 -SERLR--QDAALSVLPALLDAVRRROELNPRAPRLRLSLSDAAROVRAAGAVEYVL 140  
QY 152 ATUGPPLPQPLPGTEPA-----WAPGPAHSDFLOKMDDFWLKELCTWLMRSKADENRL 205  
DB 141 AALGAARG--PGPEPVYATLTFTANSTAGIFSAYLGFHVGLGVEWVSRTFGDGLGOL 197

RESULT 13  
US-08-443-129-3  
Sequence 3, Application US/08443129  
Patent No. 5627073

GENERAL INFORMATION:  
APPLICANT: Baker, Joffre  
APPLICANT: Chien, Kenneth  
APPLICANT: King, Kathleen  
APPLICANT: Pennica, Diane  
APPLICANT: Wood, William  
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:



